

QY 2036 GGGCCGAGGAGGGCGGGGAAGTTCAGCCAGTTTGGCAGAAACACACGAGCCCGCCTC 2095
DB |||||
695 GGGGCGGAGGAGGGCGGGGAAGTTCAGCCAGTTTGGCAGAAACACACGAGCCCGCCTC 744
QY |||||
2096 CCGCCCCCAGGGCTTCTGAGGCTTCAGCCATCCACATTCACATCCACTCGATCTCTC 2155
DB |||||
745 CCGCCCCCAGGGCTTCTGAGGCTTCAGCCATCCACATTCACATCCACTCGATCTCTC 804
QY |||||
2156 CTGAATCCCAAGACGCTATCCCTTTTCTGCTGCAATCGTCTCTGTACATGTGTGTACATAT 2275
DB |||||
805 CTGAATCCCAAGACGCTATCCCTTTTCTGCTGCAATCGTCTCTGTACATGTGTGTACATAT 864
QY |||||
2216 CAAGCAAAATGCACACCCCTTTTCTGCTGCAATCGTCTCTGTACATGTGTGTACATAT 2275
DB |||||
865 CAAGCAAAATGCACACCCCTTTTCTGCTGCAATCGTCTCTGTACATGTGTGTACATAT 924
QY |||||
2276 AGAAAGGGAAGATGTTAAGATATGTGCGCTGTGGGTTTACACAGGCTGCGTGCAGCGGTAA 2335
DB |||||
925 AGAAAGGGAAGATGTTAAGATATGTGCGCTGTGGGTTTACACAGGCTGCGTGCAGCGGTAA 984
QY |||||
2336 TATATTTTGAATAATATATCAATAACTCAACTCACTCAATTTTAAATCAATATTA 2395
DB |||||
985 TATATTTTGAATAATATATCAATAACTCAACTCACTCAATTTTAAATCAATATTA 1044
QY |||||
2396 ATTATTTTCTTTTAAAGAGAAAGCAGGCTTTCTAGACTTTTAAAGATAAAGTCTTT 2455
DB |||||
1045 ATTATTTTCTTTTAAAGAGAAAGCAGGCTTTCTAGACTTTTAAAGATAAAGTCTTT 1104
QY |||||
2456 GGGAGGCTCTACGGTGTAGAGAGGAGCTTTGAGGCCACCGCACAAAATTCACCCAGG 2515
DB |||||
1105 GGGAGGCTCTACGGTGTAGAGAGGAGCTTTGAGGCCACCGCACAAAATTCACCCAGG 1164
QY |||||
2516 GAAATCTGTCGGAAGGACACTCAGGCGAGTTCGGATCACTGTGTATGTCACAGAAG 2575
DB |||||
1165 GAAATCTGTCGGAAGGACACTCAGGCGAGTTCGGATCACTGTGTATGTCACAGAAG 1224
QY |||||
2576 GGATACGCTCTCTTGAAGAGGAACTCTGTCACTCTCTCATGCCCTGTCTAGCTCATACAC 2635
DB |||||
1225 GGATACGCTCTCTTGAAGAGGAACTCTGTCACTCTCTCATGCCCTGTCTAGCTCATACAC 1284
QY |||||
2636 CCATTTCTCTTGTCTTCACAGGTTTAACTGGTTTTTGGANATCTGTATATATATCTC 2695
DB |||||
1285 CCAATTTCTCTTGTCTTCACAGGTTTAACTGGTTTTTGGANATCTGTATATATATCTC 1344
QY |||||
2696 TGTCTCTCTCTGTATCTCTCCCTCCCTCCCTCCCTCTCTCTCCATCTCCATCTT 2755
DB |||||
1345 TGTCTCTCTCTG-TTATCTCTCCCTCCCTCCCTCCCTCCCTCTCTCTCCATCTCCATCTT 1403
QY |||||
2756 TTGAATTTCTCTCATCCCTCCATCTCAATCCCGTATCTACGCA--CCCCCCCCCCCCCAGG 2813
DB |||||
1404 TTGAATTTCTCTCATCCCTCCATCTCAATCCCGTATCTACGCAACCCGCCGCCGCCGCCAGG 1463
QY |||||
2814 CAAGCAGTGTCTCTGAGTATCATCATCACACAAAGGAACAAAGCGAACAACACACAAACA 2873
DB |||||
1464 CAAGCAGTGTCTGAGTATCATCATCACACAAAGGAACAAAGCGAACAACACACAAACA 1523
QY |||||
2874 GCCTCACTTACATCTGGTTACTCAAAAGACAGAGTCAATGGTACTTGTCTAGCGTT 2933
DB |||||
1524 GCCTCACTTACATCTGGTTACTCAAAAGACAGAGTCAATGGTACTTGTCTAGCGTT 1583
QY |||||
2934 TTGGAAGAGAAACAGGAACCCACAAACCAACCAATCAACCAACCAAGAAAAAATTC 2993
DB |||||
1584 TTGGAAGAGAAACAGGAACCCACAAACCAACCAATCAACCAACCAAGAAAAAATTC 1643
QY |||||
2994 CACAATGAAAGAAATGATTTTGTCTTTTGCATTTTGGTGTATAGCCATCAATATTCAG 3053
DB |||||
1644 CACAATGAAAGAAATGATTTTGTCTTTTGCATTTTGGTGTATAGCCATCAATATTCAG 1703
QY |||||
3054 CAAATGATCTCTCTTTT-AAAAAATAATGTGGAGGAAGTAGAAAAATTTTACCAAGGT 3112
DB |||||
1704 CAAATGATCTCTCTTTTAAAAAATAATGTGGAGGAAGTAGAAAAATTTTACCAAGGT 1763

QY 3113 TGTGGCCCGAGGGCGTTAAATTCAGATTTTTTAAACGAGAAAAACACACAGAAAGAC 3172
DB |||||
1764 TGTGGCCCGAGGGCGTTAAATTCAGATTTTTTAAACGAGAAAAACACACAGAAAGAC 1823
QY |||||
3173 TACCTCAGGTGTATTTTACCTCAGCACTTGCCTTGTCTTCCCTTAGAGATTTTGTAAA 3232
DB |||||
1824 TACCTCAGGTGTATTTTACCTCAGCACTTGCCTTGTCTTCCCTTAGAGATTTTGTAAA 1883
QY |||||
3233 GCTGATAGTTGAGCATTTTTTTTAAATAAATAAGTGTGGAAAAAATAAGA 3292
DB |||||
1884 GCTGATAGTTGAGCATTTTTTTTAAATAAATAAGTGTGGAAAAAATAAGA 1943
QY |||||
3293 TATCAACTGCCAGCCTGGAGAGGTGCAGTCCAAAGTGTGCAACAGCTTCTGAAATTTGT 3352
DB |||||
1944 TATCAACTGCCAGCCTGGAGAGGTGCAGTCCAAAGTGTGCAACAGCTTCTGAAATTTGT 2003
QY |||||
3353 CTTCGCTAGCCAAAGAACCNATATGGCCTTTCTTTGGACAAACCTTGAAAAATGTTTATTT 3412
DB |||||
2004 CTTCGCTAGCCAAAGAACCNATATGGCCTTTCTTTTGGACAAACCTTGAAAAATGTTTATTT 2063

RESULT 8
US-10-262-445-39
; Sequence 39, Application US/10262445
; Publication No. US20040014058A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John
; APPLICANT: Burgess, Catherine
; APPLICANT: Catterton, Elina
; APPLICANT: Chant, John
; APPLICANT: Chaudhuri, Amitabha
; APPLICANT: Edinger, Shlomit
; APPLICANT: Gerlach, Valerie
; APPLICANT: Glot, Loic
; APPLICANT: Gorman, Linda
; APPLICANT: Guo, Xiaojia
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Mezes, Peter
; APPLICANT: Millet, Isabelle
; APPLICANT: Ooi, Chean Eng
; APPLICANT: Patturajan, Meera
; APPLICANT: Rieger, Daniel
; APPLICANT: Spytek, Kimberly
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Zehrusen, Bryan
; APPLICANT: Zhong, Haihong
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS
; FILE REFERENCE: 21402-462D
; CURRENT APPLICATION NUMBER: US/10/262,445
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/327,454
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,849
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/329,414
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 60/330,142
; PRIOR FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: 60/341,058
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: 60/343,629
; PRIOR FILING DATE: 2001-10-24
; PRIOR APPLICATION NUMBER: 60/349,575
; PRIOR FILING DATE: 2001-10-29
; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 133
 ; SOFTWARE: CurationList version 0.1
 ; SEQ ID NO 39
 ; LENGTH: 1707
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (5)..(1669)
 US-10-262-445-39

Query Match 41.7%; Score 1423.8; DB 16; Length 1707;
 Best local similarity 88.4%; Pred. No. 0;
 Matches 1625; Conservative 0; Mismatches 82; Indels 132; Gaps 2;

QY	66	ACGGATGATGAACAGCTTTACATCGGGAACCTGAGCCCGCGCTGACCCGCGAGACCT	125
DB	1	ACGGATGATGAACAGCTTTACATCGGGAACCTGAGCCCGCGCTGACCCGCGAGACCT	60
QY	126	CCGGCAGCTCTTTGGGACAGGAGCTGCCCTGGCGGAGCAGGTCTGTGAAGTCGG	185
DB	61	CCGGCAGCTCTTTGGGACAGGAGCTGCCCTGGCGGAGCAGGTCTGTGAAGTCGG	120
QY	186	CTAGCGCTTGTGGACTACCCGACAGAACTGGGCCATCCGCGCATCGAGACCTCTC	245
DB	121	CTAGCGCTTGTGGACTACCCGACAGAACTGGGCCATCCGCGCATCGAGACCTCTC	180
QY	246	GGTAAAGTGGAAATTCATGCGGAATATGGAAGTGTGATCTCAGTCTCTATAAAGCT	305
DB	181	GGTAAAGTGGAAATTCATGCGGAATATGGAAGTGTGATCTCAGTCTCTATAAAGCT	240
QY	306	AAGGACGAGGAAATTCAGATTCGAAACATCCCTCTCCTGAGTGGAGGTGTGGA	365
DB	241	AAGGACGAGGAAATTCAGATTCGAAACATCCCTCTCCTGAGTGGAGGTGTGGA	300
QY	366	TGGAATTTGGCTCAATATGGACAGTGGAGAAATGTGGAACAGTCAACACAGACAGA	425
DB	301	TGGAATTTGGCTCAATATGGACAGTGGAGAAATGTGGAACAGTCAACACAGACAGA	360
QY	426	AACGCGCTTGTCAAGTCAATATGCAACAGAGAAAGCAAAATAGCCATGGAGAA	485
DB	361	AACGCGCTTGTCAAGTCAATATGCAACAGAGAAAGCAAAATAGCCATGGAGAA	420
QY	486	GCTAAGCGGGCATCAGTTTGAGAACTACTCTTCAAGATTTCTTACATCCCGGATGAA	545
DB	421	GCTAAGCGGGCATCAGTTTGAGAACTACTCTTCAAGATTTCTTACATCCCGGATGAA	480
QY	546	GGTGAAGTCCCTTGGCCCTCAGCGAGCCAGCGTGGGACCACTCTCCCGGAGCA	605
DB	481	GGTGAAGTCCCTTGGCCCTCAGCGAGCCAGCGTGGGACCACTCTCCCGGAGCA	540
QY	606	AGGCCAGCCCTGGGGGCACTTCTAGGCCAGACAGATTGATTTCCCGTCCGATCCT	665
DB	541	AGGCCAGCCCTGGGGGCACTTCTAGGCCAGACAGATTGATTTCCCGTCCGATCCT	600
QY	666	GGTCCCCACCCAGTTTCTGGTGCCATCATCGGAAGAGGGCTTGACCAATAAGAACAT	725
DB	601	GGTCCCCACCCAGTTTCTGGTGCCATCATCGGAAGAGGGCTTGACCAATAAGAACAT	660
QY	726	CATTAAGCAGCCAGTCCCGGTAGATATCCATAGAAAAGAGAACTCTGGAGCTGAGA	785
DB	661	CATTAAGCAGCCAGTCCCGGTAGATATCCATAGAAAAGAGAACTCTGGAGCTGAGA	720
QY	786	GAAGCTGTCAACATCCATCCCGGAGAGGAGCTTCTGAAGCATGCGCATGATTTCT	845
DB	721	GAAGCTGTCAACATCCATCCCGGAGAGGAGCTTCTGAAGCATGCGCATGATTTCT	780
QY	846	TGAAATCATGCAAGAGGCGAGATGAGACCAACTAGCCGAGAGATTCCTCTGAAAT	905
DB	781	TGAAATCATGCAAGAGGCGAGATGAGACCAACTAGCCGAGAGATTCCTCTGAAAT	840
QY	906	CTTGGCACAATGGCTTGGTGGAGACTGATTTGGAAAAGAGGAGCAAAATTTGAAGAA	965

Db	841	CTTGGCCCAACAATGGCTTGGTGGAGACTGATTTGAAAAGAGGACGAGAAATTTGAAGAA	900
QY	966	AATTGAACATGAAACAGGGACCAAGATAAACAATCTCTCTTTTGGAGGATTTGAGCATATA	1025
Db	901	AAATGAACATGAAACAGGGACCAAGATAAACAATCTCTCTTTTGGAGGATTTGAGCATATA	960
QY	1026	CNACCCGAAAGAACCAATCACTGTGAAGGGACAGTTGAGGCTGTGCGAGTGTGAGAT	1085
Db	961	CAACCCGAAAGAACCAATCACTGTGAAGGGACAGTCGAGGTCTGTGCGAGTGTGAGAT	1020
QY	1086	AGAGATTATCAAGAAGCTGCTGAGGCTTTGAAAATGATATGCTGCTGTTAAACCAACA	1145
Db	1021	AGAGATTATCAAGAAGCTGCTGAGGCTTTGAAAATGATATGCTGCTGTTAAACCAACA	1074
QY	1146	AGCAATATGATCCCAAGGTTGAACCTCAGCGCACTTGGCATCTTTTCAACAGACTGTC	1205
Db	1075	-----	1074
QY	1206	CGTGTCTATCTCCACAGCAGGGCCCGGGAGCTCCCGCGTTCGCCCTTACCACCCCTT	1265
Db	1075	-----	1074
QY	1266	CATACCCACTCCGGATCTTCTCCAGCCTGTACCCCGATCACAGTTTGGCCCGTTCCT	1325
Db	1075	---TACCCACTTCGGATCTTCTCCAGCCTGTACCCCGATCACAGTTTGGCCCGTTCCT	1131
QY	1326	GCATCATCTCTTATCCAGAGCAGGAGATTGTGAATCTTTCATCCCAACCCAGGCTGT	1385
Db	1132	GCATCATCTCTTATCCAGAGCAGGAGATTGTGAATCTTTCATCCCAACCCAGGCTGT	1191
QY	1386	GGGCGCCATCATCGGGAAGAGGGGCGACATCAAAACAGCTGGCGAGATTCGTGGAGC	1445
Db	1192	GGGCGCCATCATCGGGAAGAGGGGCGACATCAAAACAGCTGGCGAGATTCGTGGAGC	1251
QY	1446	CTTATCAAGATTGCCCTCGGGAAGGCCAGCTCAGCGAAGAGTGTGTCATCATCAC	1505
Db	1252	CTTATCAAGATTGCCCTCGGGAAGGCCAGCTCAGCGAAGAGTGTGTCATCATCAC	1308
QY	1506	CGGCGCAACCGGAAGCCAGTTTCAAGGCCCGAGGCGGATCTTTGGGAACTGAAAGAGA	1565
Db	1309	CTGGCGCAACCGGAATCCAGTTTCAAGGCCCGAGGCGGATCTTTGGGAACTGAAAGAGA	1368
QY	1566	AACTCTTTTAAACCCCAAGAGAGTGAAGTGAAGCTGGAAGCGCATATCAGAGTSCCTTTC	1625
Db	1369	AACTCTTTTAAACCCCAAGAGAGTGAAGTGAAGCTGGAAGCGCATATCAGAGTSCCTTTC	1428
QY	1626	CACAGCTGGCGGGTGTGTTGGCAAGGTGGCAAGCGCTGAACTGGAAGCTTAAAC	1685
Db	1429	CACAGCTGGCGGGTGTGTTGGCAAGGTGGCAAGCGCTGAACTGGAAGCTTAAAC	1488
QY	1686	CAGTGCAGAGTGTATCTGCTCTGTGACCAAAACCGCAGATGAAATGAGGAAGTGCCT	1745
Db	1489	CAGTGCAGAGTGTATCTGCTCTGTGACCAAAACCGCAGATGAAATGAGGAAGTGCCT	1548
QY	1746	CAGAAATTATCGGCACCTTCTTCTGCTAGCCAGACTGCAACAGCGAAGATCAGGGAATTT	1805
Db	1549	CAGAAATTATCGGCACCTTCTTCTGCTAGCCAGACTGCAACAGCGAAGATCAGGGAATTT	1608
QY	1806	ACAAAGGTTGAAGCAGCAGGAGCAGAAATACCTCTCAGGAGTGCCTCAGCGCAGCAA	1865
Db	1609	ACAAAGGTTGAAGCAGCAGGAGCAGAAATACCTCTCAGGAGTGCCTCAGCGCAGCAA	1668
QY	1866	GTGAGCTCCACAGCAGCAGCAAAACCGATGAAT	1904
Db	1669	GTGAGCTCCACAGCAGCAGCAAAACCGATGAAT	1707

RESULT 9
 US-09-764-864-749
 ; Sequence 749, Application US/09764864
 ; Patent No. US20020132753A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.

		TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies	
		FILE REFERENCE: PT223	
		CURRENT APPLICATION NUMBER: US/09/764,864	
		CURRENT FILING DATE: 2001-01-17	
		Prior application data removed - consult PALM or file wrapper	
		NUMBER OF SEQ ID NOS: 1792	
		SOFTWARE: PatentIn Ver. 2.0	
		SEQ ID NO 749	
		LENGTH: 1186	
		TYPE: DNA	
		ORGANISM: Homo sapiens	
		FEATURE:	
		NAME/KEY: SITE	
		LOCATION: (511)	
		OTHER INFORMATION: n equals a,t,g, or c	
		NAME/KEY: SITE	
		LOCATION: (774)	
		OTHER INFORMATION: n equals a,t,g, or c	
		NAME/KEY: SITE	
		LOCATION: (777)	
		OTHER INFORMATION: n equals a,t,g, or c	
		US-09-764-864-749	
		Query Match 32.1%; Score 1096.8; DB 9; Length 1186;	
		Best Local Similarity 98.4%; Pred. No. 1.4e-290;	
		Matches 1145; Conservative 3; Mismatches 12; Indels 4; Gaps 4;	
QY	1376	CCGAGCTGTGGCGCCATCATCGGGAAGAGGGGCGACACATCAAAACAGCTGGCGAGAT	1435
DB	25	CCCAAGCTGTGGCGCCATCATCGGGAAGAGGGGCGACACATCAAAACAGCTGGCGAGAT	84
QY	1436	TGCGCGGAGCTCTATCAAGATGTCCTCGGGAAGGCCAGAGCTCAGCGAAAGGATGG	1495
DB	85	TGCGCGGAGCTCTATCAAGATGTCCTCGGGAAGGCCAGAGCTCAGCGAAAGGATGG	144
QY	1496	TCATCATCACCGGCGCCACCGGAAGCCAGTTCAGGCGCCAGGACGGATCTTTGGGAAC	1555
DB	145	TCATCATCACCGGCGCCACCGGAAGCCAGTTCAGGCGCCAGGACGGATCTTTGGGAAC	203
QY	1556	TGAAGAGGAAACTTCTTAAACCCAAAGAGAGTGAAGCTGGAAGCGCATATCAGAG	1615
DB	204	TGAAGAGGAAACTTCTTAAACCCAAAGAGAGTGAAGCTGGAAGCGCATATCAGAG	263
QY	1616	TGCGCTCTTCCACAGCTGCGCGGTGATTTGGCAAGGTGGCAAGCCGTTGAACGAATGC	1675
DB	264	TGCGCTCTTCCACAGCTGCGCGGTGATTTGGCAAGGTGGCAAGCCGTTGAACGAATGC	323
QY	1676	AGAACTTAAACAGTGAGAGTCACTGCTGTCGACCAACCGCCAGATGAATGAGG	1735
DB	324	AGAACTTAAACAGTGAGAGTCACTGCTGCTCGTGACCAACCGCCAGATGAATGAGG	383
QY	1736	AGTGTATCTCGAATATTCGGGCACTTTCTTGTAGCCAGATGACACGCGCAAGATCA	1795
DB	384	AGTGTATCTCGAATATTCGGGCACTTTCTTGTAGCCAGATGACACGCGCAAGATCA	443
QY	1796	GGGAATTTG-TACAAAGGTGAAGCAGCAGGAGCAGAAATACCTCAGGAGTCGCTCA	1854
DB	444	GGGAATTTG-TACAAAGGTGAAGCAGCAGGAGCAGAAATACCTCAGGAGTCGCTCA	503
QY	1855	CAGCGAGCAAGTGAAGTCTCCACAGGCCACAGCAACCGAGATGAATGAGCCCTTC	1914
DB	504	CAGCGCA-NAAGTGAGGCTCCACAGGCCACAGCAACCGAGATGAATGAGCCCTTC	562
QY	1915	CAACCTGACAGATCAGACCAACCGCCAGCAGATCGGAGCAAAACCAAGACCA	1974
DB	563	CAACCTGACAGATCAGACCAACCGCCAGCAGATCGGAGCAAAACCAAGACCA	622
QY	1975	TCTGAGCAATGAGAAGTCTGCGGAGGCGCCAGGAGCTCTGCCAGGCCCTGAGAACCC	2034
DB	623	TCTGAGCAATGAGAAGTCTGCGGAGGCGCCAGGAGCTCTGCCAGGCCCTGAGAACCC	682
QY	2035	AGGGCGCAGGAGGGGGGAAGGTGAGCCAGGTTTGCCAGAACCCAGCCCGCCCT	2094
		US-09-764-864-330	
		Query Match 22.7%; Score 776.2; DB 9; Length 822;	
		Best Local Similarity 97.8%; Pred. No. 2.1e-202;	
		Matches 805; Conservative 10; Mismatches 5; Indels 3; Gaps 3;	
QY	779	CTCAGAGAAGCTGTCACTCCATCCAGCCGCCAGAGGGGACTTCTGAGCATCGCGCA	838
DB	1	CTCAGAGAAGCTGTCACTCCATCCATCCATCCATCCATCCATCCATCCATCCATCCAT	60
QY	839	TGATTTCTTGAATCATCGAGAAAGAGCAGATCAGACCAAACTAGCCGAGAGATTCCTC	898
DB	61	TGATTTCTTGAATCATCGAGAAAGAGCAGATCAGACCAAACTAGCCGAGAGATTCCTC	120


```
1024 TACAAACCCGGAAGAACCACTCACTGTGAAGGGCACAGTTGAGCCCTGTGCCAGTGTGAG 1083
Db      |||
961 TATATCCAGAGCCACTATTACAGTTAAAGGCAATGTTGAGACATGTGCCAAAGCTGAG 1020
QY      |||
1084 ATAGAGATTATGAAGAGCTGCTGAGGCTTTTGAATAATGATATGCTGTTTACCAA 1143
Db      |||
1021 GAGGAGATCATGAAGAAATCAGGAGTCTTATGAATGATATGCTTCTATGAATCTT 1080
QY      |||
1144 CAGCCCAATCTGATCCAGGGTTGAACCTCAGCGCACTTGGCATCTTTTCAACAGGACTG 1203
Db      |||
1081 CAAGCACATTAATCTCGATTAATCTGAAGCTTGGCTCTGTTCCACCCACTTCA 1140
QY      |||
1204 TCCGTGCTATCTCACACAGCAGGCGCCCGCGAGCTCCCCCGCTGCCCCCTACACCCC 1263
Db      |||
1141 GGGATGCCACTCCCACTCAGGGCCCCCTTCAGCCATGACTCT 1185
QY      |||
1264 TTCACTACCACTCCCGATATCTTCAGCGTGTATCCCGCTGTACCCCATCACAGTTTGGCCGTTT 1323
Db      |||
1186 -----CCCTACCGCGAGTTTGAGC----- 1204
QY      |||
1324 CCGCATCATCACTCTTATCCAGACGAGAGATTGTGAATCTCTTCACTCCCAACCCAGGCT 1383
Db      |||
1205 -----AATCAGAAACGAGACTGTTCTATCTGTTTATCCAGCTCTATCA 1248
QY      |||
1384 GTGGCGCCATCATCGGGAAGAGGGGGCACATCAAAACAGCTGGCGAGATTCCCGGA 1443
Db      |||
1249 GTCGTGCCATCATCGGACAGCAGGGCCAGCACATCAAGCAGCTTTCTCGCTTTGCTGA 1308
QY      |||
1444 GCCTCTATCAAGATTGCCCTGCGGAGGCCAGACGTCACGAAAGGATGTCATCATC 1503
Db      |||
1309 GCTTCAATTAAGATTGCTCCAGCGGAAGCACACAGATGCTAAAGTGAGGATGTTGATATC 1368
QY      |||
1504 ACCGGCCACCGGAGCCAGTTCAAGCCCGAGGACGGATCTTTGGGAAACTGAAAGAG 1563
Db      |||
1369 ACTGGACCACGAGGCTCAGTTCAAGGCTCAGGGAAGATTTAGGAAATTAAGAA 1428
QY      |||
1564 GAAACTTCTTTAAACCCCAAGAGAGTGAAGCTTGAAGCGCATPATCAGAGTGCCTCT 1623
Db      |||
1429 GAAACTTTGTAGTCTTAAGAGAGGTGAACCTTGAAGCTCATPATCAGAGTGCCATCC 1488
QY      |||
1624 TCCACAGTGGCCGGTGTATGGCAAGGTGGCAAGACGGTGAACGAACACTGAGACTTA 1683
Db      |||
1489 TTTGCTGTGCAGAGTTATTGGAAAGGAGGCAAAACGGTGAATGAACCTCAGAAATTG 1548
QY      |||
1684 ACCAGTCAGAGTCACTGTGCTCGTGACCAACCGCAGATGAAATGAGGAAGTGAATC 1743
Db      |||
1549 TCAAGTGCAGAGTTGTTGTCCTGTGTACAGACACCTGTATGAGATGAGCAAGTGGT 1608
QY      |||
1744 GTCAGAAATTATCGGCACTTTCTTGTAGCCAGCTGCAAGCGCAAGATCAGGAAATT 1803
Db      |||
1609 GTCAAATTAACCTGGTCACTTCTATGTTGCCAGTTGCCCAGAGAAATTCAGGAAATT 1668
QY      |||
1804 GTACACAGGTGAAGCAGAGCAGAAATACCTTCAGGAGTCGCTCAC 1855
Db      |||
1669 CTGACTCAGGTAAAGCAGCAACCAACAGAAAGGCTCTCCTCAAGTGGACAC 1720
```

RESULT 12

```
US-09-850-716A-347
; Sequence 347, Application US/09850716A
; Patent No. US20020115139A1
; GENERAL INFORMATION:
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Retter, Marc W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C15
; CURRENT APPLICATION NUMBER: US/09/850,716A
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 3.0
```

```
; SEQ ID NO 347
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-850-716A-347
```

```
Query Match      20.9%; Score 712; DB 9; Length 1740;
Best Local Similarity 64.5%; Pred. No. 1.8e-184;
Matches 1156; Conservative 0; Mismatches 555; Indels 81; Gaps 3;
```

```
QY      73 ATGAACAAGCTTTACATCGGGAACCTGTAGCCCGCGCTCACCGCCGACGACCTCCGCGAG 132
Db      |||
1 ATGAACAATCTGATATCGGAACCTCAGCGAGAAGCGCGCCCTCGGACCTTAGAAGT 60
QY      133 CTCTTTGGGACAGGAGCTGCCCTTGGGGGACAGTCTCTGCTGAAGTCCGGCTACGCC 192
Db      |||
61 ATCTTCAAGAGACCCCAAGATCCCGGTGTCGGGACCCCTTCTGTGAAGACTGCTACGCG 120
QY      193 TTCGTGACTACCCCGACAGAACTCGGGCCATCCGGCCATCGAGACCCCTTCTGGGTAAA 252
Db      |||
121 TTCTGTGACTGCCCGGACGAGAGCTGGGCCCTCAAGGCCATCGAGGGCTTTTCAGGTAAA 180
QY      253 GTGGAATTCATGGGAAAAATCATGGAAGTTGATCTCAGTCTCTAAAGACTTAAGAGC 312
Db      |||
181 ATAGAACTGCACGGGAAACCCATAGAAATTTGAGCACTCGGTCCCAAAAAGGCAAGGATT 240
QY      313 AGGAAATTTCAGATTCGAAACATCCCTCACCTGCAGTGGGAGGTGTTGGATGGAATT 372
Db      |||
241 CGGAAATCTCAGATACGAAATATCCCCCTCATTTACAGTGGGAGGTGCTGGATAGTTA 300
QY      373 TTGGCTCAATATGGAACAGTGGGAATGTGGAAACAAGTCAACACAGACACAGAAACCGCC 432
Db      |||
301 CTAGTCCAGTATGAGTGGTGGAGAGCTGTGAGCAAGTGAACACTGACTCGCGAAACTGCA 360
QY      433 GTTGTCAACGTCAATATGCAACAGAGAGAGCAAAAATAGCCATGGAAGACTAAGC 492
Db      |||
361 GTTGTAAATGTAACTATTTCAGTAAGACCAAGCTAGACAAGCACTAGACAACTGAAT 420
QY      493 GGGCATCAGTTTGAGAACTACTCTTCAAGATTTCCTACATCCCGGATGGAAGGTGAGC 552
Db      |||
421 GGAATTCAGTTAGAGAAATTCACCTTGAAAGTAGCCTATATCCCTGATGAAACGCGCCGC 480
QY      553 TCCCTTTCGCCCCCTCAGCGAGCCCGAGCTGGGGACCACTCTTCCCGGGAGCAAGGC--- 609
Db      |||
481 CAGCAAAACCCCTTCAGCAGACCCCGAGGTTCGCGGGGCTTGGGCAGAGGGGCTCTCTCA 540
QY      610 -----CAGGCCCTTGGGGGCACTTCTCAGGCCAGACAGATGATTTCCCGTGGGATC 663
Db      |||
541 AGGCAAGGGGTCTCCAGGATCCGTATCCAAAGCAGAAACCACTGTGATTTGCTCTGCGCCTG 600
QY      664 CTGGTCCCAACCCAGTTTGTGGTCCCATCATCGAAAGAGGGGCTTGACCAATAAGAAC 723
Db      |||
601 CTGGTTCACCAACCAATTTGTTGAGCCCATCATAGAAAGAGGTGCCACCATTCGGAAC 660
QY      724 ATCACTAAGCAGACCCAGTCCCGGGTAGATATCCATAGAAAAAGAGAACTCTCGAGTGA 783
Db      |||
661 ATCAACCAACAGACCCAGTCTAAATCGATGTCCACCGTAAAGAAAATCGCGGGGCTGCT 720
QY      784 GAGAGCCCTGTCCCAATCCATGCCACCCAGAGGGGACTCTTGAAGCATGCCGCATGATT 843
Db      |||
721 GAGAGTCGATTAATCTCTCTACTCTCTGAAAGCACCCTCTGCGGCTTTGTAAGTCTATT 780
QY      844 CTTGAAATCATGCAGAAAGAGCCAGATGAGACAAACTAGCCGAGAGATTCCTCTGAAA 903
Db      |||
781 CTGGAGATTATGCATAAGGAAGCTCAAGATATAAATTCACAGAGAGATCCCTTTGAAG 840
QY      904 ATCTTGGCACACAAATCGCTTGGTTGGAAGACTGATTGGAAGAAAGAGCGGAAATTTGAAG 963
Db      |||
841 ATTTTACCTCATATTAACCTTTGTTGACGCTTATTGTTGTAAGAGGAAAGAAATCTTAAA 900
QY      964 AAAATTGAAATGAAAACAGGACCAAGATACAACTCATCTTTGAGGATTTGAGCAT 1023
Db      |||
901 AAAATTGAGCAAGACACAGACACATAAATCACGATATCTCCATTTGAGGAATTTGACGCTG 960
```


[illegible]

RESULT 14

RESULT 14
US-10-007-700-347

US-10-007-700-347 : Sequence 347. Application US/10007700

; Sequence 347, Application US/10

; Publication No. US20

GENERAL INFORMATION:

APPLICANT: Wang, Tongtong

APPLICANT: Wang, Aijun

APPLICANT: Skeiky, Yasyr

```

; APPLICANT: Kalos, Michael D.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; APPLICANT: Retter, Marc W.
; APPLICANT: Durham, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Peckman, David W.
; APPLICANT: Cai, Feng
; APPLICANT: Foy, Teresa M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C17
; CURRENT APPLICATION NUMBER: US/10/007,700
; CURRENT FILING DATE: 2001-11-30
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 347
; LENGTH: 1740
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-007-700-347

```

Query Match 20.9%; Score 712; DB 13; Length 1740;

Query Match	Score	Score	Score	Score
20.5%	20.5%	20.5%	20.5%	20.5%
Best Local Similarity	64.5%	Pred. NO. 1.8e-184		
Matches 1156:	Conservative	0;	Mismatches 555;	Indels 81;
Gaps	3;			

QY	73	ATGAACAAGCTTTACATCGGGAACCTGAGCCCGCGCTCAACGCGGACAGCCTCCGGCAG	133
DB	1	ATGAACAACCTGTATATCGGAATCTCAGCGAAGCCGCGCCCTCGGACTAGAAGT	60
QY	133	CTCTTTGGGGACAGGAAGCTCCGCTCGCGGGACAGGTCCTGTGAAGTCCGGTACGCC	192
DB	61	ATCTTCAAGGACGCCAAGATCCCGGTGTCGGGACCTTCTCTGTGAAGACTGGCTACGCG	120
QY	193	TTCTGTGGACTACCCGGACCGAATCTGGGCGATTCGCGGCCATCGAGACCTCTCGGTTAAA	252
DB	121	TTCTGTGGACTCGCCGGACGAGAGCTGGGCGCCTCAAGGCCATCGAAGCGCTTTTCAAGTTAAA	180
QY	253	GTGGGAATTCGATGGGAAAATCATGGAAGTTGATTACTCAGTCTCTTAAAAGACTAAGGAGC	312
DB	181	ATAGAACTGCACGGGAAAACCCATGAAGTTGAGCACTCGTCCCAAAGGCAAGGATT	240
QY	313	AGGAAAATTCAGATTCGAAAATCCCTCCTCACTCGAGTGGGAGGTGTTGATGGACTTT	372
DB	241	CGGAAACTTCAGATACGAAATATCCGCGCTCATTTACAGTGGGAGGTGCTGGATAGTTTA	300
QY	373	TTGGGCTCAATATGGGACAGTGGAGAATGTGNAACAGTTCACACAGACACAGAAACCGCC	432
DB	301	CTAGTCCAGTATGGAGTGTGGAGAGCTGTGAGCAAGTGAACACTGACTCGGAAACTGCA	360
QY	433	GTGTGCAAGCTCACATATGCAACAAGAGAAGCAAAAATAGCCATGGAGAAGCTAAGC	492
DB	361	GTGTGTAATGTAACTATTCAGTAGGACCAGCTAGACAAGCACTAGACAAACTGAAT	420
QY	493	GGGCATCAGTTTGAGAACTACTCTTTCAAGATTTCCGTACATCCCGATGAAGAGTGAGC	552
DB	421	GGATTTTCAGTTAGAGAAATTCACCTTTGAAGTAGCTATATCCCTGATGAACCGCCGCCG	480
QY	553	TCCCTTTCCGCGCCCTCAGCGAGCCGAGCTGTGGGACCACCTCTTCCCGGAGCAAGCC	609
DB	481	CAGCAAAAACCCCTTGACAGAGCCCGAGGTCGCGGGGGCTTTGGGAGAGGGGCTCTCA	540
QY	610	-----CAGCGCCCTCGGGGCACTTCTCAGGGCCAGACAGATTTGATTTCCCGCTGGGATC	663
DB	541	AGCAGGGGTCTCCAGATCCGTATCCAAAGCAGAAACCATGTGATTTTGCTCTGCGCCTG	600
QY	664	CTGGTCCCAACCCAGTTTGTGTGGTGCATCATCGGAAAGAGGGCTTGACCATAAAGAC	723
DB	601	CTGGTTCCCAACCCAAATTTGTGTGGAGCCATCATAGGAAAGAGAGGTGCCATATCGGAAC	660

QY	610	-----CAGCCCTCGGGGCACTTCTCAGGCCAGACAGATTGATTTCCGCTCGGGATC	663
Db	541	AGCAGGGGTCTCCAGGATCCGATATCAAGCAGAAACCATGTGATTTGCTCTGCGCCTG	600
QY	664	CTGGTCCCAACCCAGTTTGTGTGTCATATCGGAAAGAGGGCTTGACCAATAAGAAC	723
Db	601	CTGGTCCCAACCCAAATTTGTGTGAGGCATCATAGGAAAGAGGTGCCACCATTCGGAAC	660
QY	724	ATCATTAGCAGACCCAGTCCCGGTAGATATCCATAGAAAGAGAACTCTGAGCTGCA	783
Db	661	ATCACCAGCAGACCCAGCTCTAAATCGATGCCACCGTAAGAAATGCGGGGCTGCT	720
QY	784	GAGAAGCTCTCACCATCCATGCCACCCAGGGGACTTCTGAAGCATGCCGATGATT	843
Db	721	GAGAGTCGATTACTTCTCTCTACTCTCTGAAAGCACCTCTGCGGCTTGTAAATCTATT	780
QY	844	CTTGAATCATGAGAAAGGAGCAGATGACACCAACTAGCCGAGAGATTCCTCTGAAA	903
Db	781	CTGGAGATTATGCATAGGAAGCTCAAGATATAAAATTCAGAAAGAGATCCCTTGAAG	840
QY	904	ATCTTGGCACACATGCTTGTGTGGAAGCTGATTTGMAAAGAAAGGCAGAAATTTGAAG	963
Db	841	ATTTTAGCTCATATAAACTTTGTGGAGCTTTATGTTAAAGAGAAATTCATTTAA	900
QY	964	AAAATTGAACATGAACAGGGACCAAGATAACAATCTCATCTTTGAGGATTTGAGCAT	1023
Db	901	AAAATTGAGCAAGACACAGACACTAAAATCAGGATATCTCCATTCAGGAATTGACGTG	960
QY	1024	TACAACCGGAAAGAACCATCAGTGAAGGACAGTGGGGCTGTGCCAGTGTGAG	1083
Db	961	TATTAATCCAGAACGCATATTACAGTTAAAGGCAATGTTGAGCATGTGCCAAAGCTGAG	1020
QY	1084	ATAGAGATTATGAAGAGCTCGCTGAGGCCCTTGAATAATGATATGCTGGCTGTGTTAACCA	1143
Db	1021	GAGGAGATCATGAGAAATCAGGAGTCTTATGAATGATTTGCTTCTATGAATCTT	1080
QY	1144	CAAGCCATCTGATCCCAAGGTTGAACCTCAGCGCATCTGGCATCTTTTCAACAGGACTG	1203
Db	1081	CAAGCACATTTAAATTCCTGGATTAAATCTGAACGCTTGGGTCTGTTCGCCACACTTCA	1140
QY	1204	TCCGTGTATCTCAGCAGAGGCGCCGCGGAGCTCCCGCGTCCCGCTACCAACCC	1263
Db	1141	GGGATGCCACCTCCACCTCAGGGCCCTTCAGCCATGACTCCT-----	1185
QY	1264	TTCACTACCCACTCCGGATCTTCTCCAGCCTGTACCCCATCACAGTTTGGCCCGTTC	1323
Db	1186	-----CCCTACCGCAGTTTGAGC-----	1204
QY	1324	CGGCATCATCTTTATCCAGAGGAGATGTGAATCTCTTATCCCAACCCAGGT	1383
Db	1205	-----AATCAGAAACGGAGACTGTTCATCTGTTTATCCAGCTCTATCA	1248
QY	1384	GTGGGGCCCATCATCGGAGAGAGGGGCACACATCAACAGCTGGCGAGATTGCGCGGA	1443
Db	1249	GTGCGTGCCCATCATCGCAAGAGGGCCAGCATCAAGCAGCTTTCTCGCTTGTCTGA	1308
QY	1444	GCTCTATCAGATTGCCCTCGGAAAGGCCAGAGCTCAGCGAAAGAGATGTCATATC	1503
Db	1309	GCTTCATTAAGATTGCTCCAGCGAGACACCATGCTTAAAGTAGGATGTTGATTATC	1368
QY	1504	ACCGGCCACCGGAAGCCAGTTTCAAGGCCAGGGACGGATCTTTGGGAACTGAAGAG	1563
Db	1369	ACTGGACCCACAGAGCTCAGTTCAAGGCTCAGGGAAGAAATTTATGGAATAATAAGAA	1428
QY	1564	GAAACTCTTTAAACCCAAAGAGAGTGAAGCTGGAAGCGCATATCAGAGTGCCTCT	1623
Db	1429	GAAACTCTTTAGTCTCTAAAGAGAGGTGAAACTTTGAAGCTCATATCAGAGTCCATCC	1488
QY	1624	TCCACAGCTGGCGGGTATTGGCAAGGTGGCAAGACCGTGAACCAACTGCAGAACTTA	1683
Db	1489	TTTGCTGCTGGCAGAGTTATTGGAAAGGAGGCAACCGTGAATGAATTCAGAAATTG	1548

Search completed: July 21, 2004, 13:09:48
Job time : 1332 secs

QY	1684	ACCAAGTCAGAAAGTCACTGCTGCTGACCAAAACGCCAGATGAAATCAGGAAGTGATC	1743
Db	1549	TCAGTGCAGAAAGTTGTTGCTCCCTGACCAAGACACCTGATGAGAATGACCAAGTGGTT	1608
QY	1744	GTCAAGATTATCGGGCACTTCTTTGCTAGCCAGACTGCAACAGGCAAGATCAGGGAAT	1803
Db	1609	GTCAAAATAACTGTGTCACTTCTATGCTGCCAGGTTGCCCAAGAAAAATTCAGGAAAT	1668
QY	1804	GTACAACAGGTGAAGCAGCAGGAGCAGAAATACCTTCAGGGAGTGCCTCAC	1855
Db	1669	CTGACTCAGGTAAGCAGCAGCAACCAACAGAGGCTCTGCAAGTGGACCAC	1720

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: July 16, 2004, 11:17:53 ; Search time 53.5 Seconds
(without alignments)
12269.364 Million cell updates/sec

Title: US-09-270-437D-6
Perfect score: 1120
Sequence: 1 ggcagcgaggaggcgagga.....aaccttgaaatgtttattt 3412

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283366 seqs, 96191526 residues
Word size: 1

Total number of hits satisfying chosen parameters: 565818

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2.1/USPTO.spool/p/US09270437/runat_16072004.113128.13969/app_query.fasta_1.3591
-DB=PIR 78 -QFMT=fastan -SUFFIX=oligo.rpr -MINMARCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09270437@cgn 1.86 @runat_16072004.113128.13969 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=60 -FGAPEXT=7
-XGAPOP=60 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : PIR 78 : *
1: pir1 : *
2: pir2 : *
3: pir3 : *
4: pir4 : *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	9	0.8	319	2 F75420	hypothetical prote
2	9	0.8	366	2 AB1422	probable GTP-bind
3	9	0.8	366	2 AB1796	probable GTP-bind
4	9	0.8	398	1 WZBN3	proteins-serine/thr
5	9	0.8	399	2 JCT957	sex-determining re
C 6	9	0.8	431	2 T29850	hypothetical prote
C 7	9	0.8	491	1 FGHUB	fibrinogen beta ch
C 8	9	0.8	696	2 A28635	transcription fact
C 9	9	0.8	788	2 JS0747	regulatory protein
C 10	9	0.8	1226	2 T45915	pre-mRNA splicing
C 11	9	0.8	1234	2 S52099	phospholipase C be
12	9	0.8	2338	2 I73957	kinase-related pro
13	9	0.8	2347	1 TVHURS	kinase-related pro
C 14	8	0.7	63	2 T12121	NADH dehydrogenase

C 15	8	0.7	121	2 F72580	hypothetical prote
C 16	8	0.7	125	2 A71245	hypothetical prote
C 17	8	0.7	137	2 S37353	modulin (clone GmE
C 18	8	0.7	137	2 G75471	hypothetical prote
C 19	8	0.7	146	2 T16341	hypothetical prote
C 20	8	0.7	150	2 T17206	hypothetical prote
C 21	8	0.7	157	2 AE1293	shikimate kinase h
C 22	8	0.7	161	1 S76604	hypothetical prote
C 23	8	0.7	171	2 B89975	conserved hypothet
C 24	8	0.7	183	2 AF1883	hypothetical prote
C 25	8	0.7	202	2 T46586	ribosomal protein
C 26	8	0.7	205	2 T34724	probable membrane
C 27	8	0.7	208	2 T16953	hypothetical prote
C 28	8	0.7	219	2 A99194	iron (III) ABC tra
C 29	8	0.7	220	2 A36298	proline-rich prote
C 30	8	0.7	242	2 G96994	glycerol uptake fa
C 31	8	0.7	257	2 C96965	transcription regu
C 32	8	0.7	267	2 D83113	probable permease
C 33	8	0.7	270	2 AF1815	superoxide dismuta
C 34	8	0.7	279	2 D82281	ferric vibriobacti
C 35	8	0.7	285	2 H70781	hypothetical prote
C 36	8	0.7	301	2 G83182	hypothetical prote
C 37	8	0.7	305	2 T20906	interferon respons
C 38	8	0.7	311	2 A38558	hypothetical prote
C 39	8	0.7	315	2 A84634	transaldolase - He
C 40	8	0.7	316	2 E71812	hypothetical prote
C 41	8	0.7	320	2 AF1892	[NiFe] uptake hydr
C 42	8	0.7	324	2 G90896	probable transcrip
C 43	8	0.7	326	2 S56534	hypothetical 36.9k
C 44	8	0.7	326	2 D91287	hypothetical prote
C 45	8	0.7	326	2 G86128	hypothetical prote

ALIGNMENTS

RESULT 1

F75420
Hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C:Accession: F75420
R:White, O.; Eissen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; S.; Smith, H.O.; Venter, J.C.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; M. Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: F75420
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-319 <WHI>
A:Cross-references: GB:AE001971; GB:AE000513; NID:G6458972; PIDN:AAF10810.1; PID:G64589
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR1229
A:Map position: 1

Alignment Scores:

Pred. No.:	11.7	Length:	319
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.82%	Indels:	0
DB:	2	Gaps:	0

US-09-270-437D-6 (1-3412) x F75420 (1-319)

QY 1079 GCATGGCACAGGCTCAACTGTGCC 1053

Db 7 AlaLeuAlaGlnAlaSerThrValPro 15

RESULT 2
AB1422

probable GTP-binding protein homolog lmo2779 [imported] - Listeria monocytogenes (strain C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AB1422
A;Status: Preliminary
A;Molecule type: DNA
A;Residues: 1-366 <GLA>
A;Cross-references: GB:NC_003210; PIDN:CAD00992.1; PID:g16412279; GSPDB:GN00177
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lmo2779
C;Superfamily: yeast probable purine nucleotide-binding protein YBR025c

Alignment Scores:
Pred. No.: 11.5 Length: 366
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
Gaps: 0
DB:

US-09-270-437D-6 (1-3412) x AB1422 (1-366)

QY 1099 AAGCTGCGTGGAGCGCTTTGAAATGAT 1125
DB 167 LysLeuArgGluAlaPheGluAsnAsp 175

RESULT 3
AH1796
Probable GTP-binding protein homolog lin2919 [imported] - Listeria innocua (strain Cliph C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AH1796
A;Status: Preliminary
A;Molecule type: DNA
A;Residues: 1-366 <GLA>
A;Cross-references: GB:AL592022; PIDN:CA98144.1; PID:g16415460; GSPDB:GN00178
A;Experimental source: strain Cliph1262
C;Genetics:
A;Gene: lin2919
C;Superfamily: yeast probable purine nucleotide-binding protein YBR025c

Alignment Scores:
Pred. No.: 11.5 Length: 366
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
Gaps: 0
DB:

US-09-270-437D-6 (1-3412) x AH1796 (1-366)

QY 1099 AAGCTGCGTGGAGCGCTTTGAAATGAT 1125
DB 167 LysLeuArgGluAlaPheGluAsnAsp 175

RESULT 4
WZBEN3
protein-serine/threonine kinase (EC 2.7.1.1-) - suid herpesvirus 1 (strain NIA-3)
N;Alternate names: UL13 protein
C;Species: suid herpesvirus 1
C;Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 11-Jun-1999
C;Accession: B42744
R;de Wind, N.; Domen, J.; Berns, A.
J. Virol. 66, 5200-5209, 1992
A;Title: Herpesviruses encode an unusual protein-serine/threonine kinase which is nones A;Reference number: A42744; MUID:92365105; PMID:1323689
A;Accession: B42744
A;Molecule type: DNA
A;Residues: 1-398 <DEW>
A;Cross-references: GB:M94870; NID:g334092; PIDN:AAA47481.1; PID:g334094
C;Genetics:
A;Gene: UL13
C;Superfamily: herpesvirus protein-serine/threonine kinase; protein kinase homology C;Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
F;78-378/Domain: protein kinase homology <KIN>
F;86-93/Region: protein kinase ATP-binding motif
F;103/Active site: Lys #status predicted

Alignment Scores:
Pred. No.: 11.4 Length: 398
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
Gaps: 0
DB:

US-09-270-437D-6 (1-3412) x WZBEN3 (1-398)

QY 2 GCAGCGGAGGCGGAGGAGCGCGGG 28
DB 119 AlaAlaGluGluAlaArgSerAlaGly 127

RESULT 5
JC7957
sex-determining region Y (SRX)-like HMG-box protein 17, SOX17 protein - rice field eel C;Species: Monopterus albus (rice field eel)
C;Date: 22-Jun-2003 #sequence_revision 22-Jun-2003 #text_change 07-Jul-2003
C;Accession: JCV957
R;Wang, R.; Cheng, H.; Xia, L.; Guo, Y.; Huang, X.; Zhou, R.
Biochem. Biophys. Res. Commun. 303, 452-457, 2003
A;Title: Molecular cloning and expression of Sox17 in gonads during sex reversal in the A;Reference number: JCV957; MUID:22546426; PMID:12659838
A;Accession: JCV957
A;Molecule type: mRNA
A;Residues: 1-399 <WAN>
A;Cross-references: GB:AY100695
C;Comment: This protein functions as a transcriptional activator during spermatogenesis.
C;Genetics:
A;Gene: Sox17
A;Map position: 5
C;Keywords: gonad differentiation; HMG box; sexual differentiation; Sox17; spermatogen

Alignment Scores:
Pred. No.: 11.4 Length: 399
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
Gaps: 0
DB:

US-09-270-437D-6 (1-3412) x JCV957 (1-399)

QY 1265 TCACATCCCATCCGGATACCTTCCCA 1291
DB 210 SerLeuProThrProAspThrSerPro 218

T29850
hypothetical protein C49C8.5 - *Caenorhabditis elegans*
C/Species: *Caenorhabditis elegans*
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C/Accession: T29850
R/Johnson, D.; Bradshaw, H.
submitted to the EMBL Data Library, June 1996
A/Description: The sequence of *C. elegans* cosmid C49C8.
A/Reference number: Z20698
A/Accession: T29850
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-431 <JOH>
A/Cross-references: EMBL:U61945; PIDN:AA03126.1; GSPDB:GN00022; CBSP:C49C8.5
A/Experimental source: strain Bristol N2; clone C49C8
C/Genetics:
A/Gene: CBSP:C49C8.5
A/Map position: 4
A/Introns: 17/3; 66/2; 106/3; 151/2; 187/1; 233/2; 302/3; 329/1; 351/2; 377/2

Alignment Scores:
Pred. No.: 11.3 Length: 431
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.82% Indels: 0
DB: 2 Gaps: 0

US-09-270-437D-6 (1-3412) x T29850 (1-431)

QY 2590 AAGGAGACGGTATCCCTTCGTTCACA 2564
|||||
Db 3 LysGluThrValSerLeuLeuLeuThr 11

RESULT 7
FGHUB
fibrinogen beta chain precursor [validated] - human
N/Alternate names: coagulation factor I
N/Contents: fibrinopeptide B
C/Species: *Homo sapiens* (man)
C/Date: 24-Apr-1984 #sequence_revision 31-Mar-1993 #text_change 08-Dec-2000
C/Accession: B43568; A90469; B90469; I37389; A94433; A90437; A94309; G54223; A03121; B37
R/Chung, D.W.; Harris, J.E.; Davie, E.W.
Adv. Exp. Med. Biol. 281, 39-49, 1990
A/Title: Nucleotide sequences of the three genes coding for human fibrinogen.
A/Reference number: A90469; MUID:91344740; PMID:2102623
A/Accession: B43568
A/Molecule type: DNA
A/Residues: 9-191, 'A', 193-491 <CHU>
R/Chung, D.W.; Que, B.G.; Rixon, M.W.; Mace Jr., M.; Davie, E.W.
Biochemistry 22, 3244-3250, 1983
A/Title: Characterization of complementary deoxyribonucleic acid and genomic deoxyribonu
A/Reference number: A90469; MUID:83283433; PMID:6688356
A/Accession: A90469
A/Molecule type: DNA
A/Residues: 1-38 <CHU>
A/Accession: B90469
A/Molecule type: mRNA
A/Residues: 9-191, 'A', 193-491 <CH2>
A/Cross-references: GB:J003129; NID:g182429; PIDN:AA52429.1; PID:g182430
R/Huber, P.; Dalmont, J.; Courtois, G.; Laurent, M.; Assouline, Z.; Marguerie, G.
Nucleic Acids Res. 15, 1615-1625, 1987
A/Title: Characterization of the 5'-flanking region for the human fibrinogen beta gene.
A/Reference number: I37389; MUID:87146483; PMID:3029722
A/Accession: I37389
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-38 <HUB>
A/Cross-references: EMBL:X05018; NID:g31400; PIDN:CAA28674.1; PID:g31401
R/Henschen, A.; Lottspeich, F.; Southan, C.; Topfer-Petersen, E.
in Protides of the Biological Fluids, Proc. 28th Colloq., Peeters, H., ed., pp.51-56, Pe
A/Title: Human fibrinogen: sequence, sulfur bridges, glycosylation and some structural v

A;Description: fibrinogen cleaved by thrombin yields monomers that are polymerized into
 A;Pathway: blood coagulation
 C;Superfamily: fibrinogen beta chain; fibrinogen beta/gamma homology; fibrinogen disulfide
 C;Keywords: blood coagulation; coiled coil; glycoprotein; liver; plasma; pyroglutamic acid
 F;1-30/Domain: (or 4-30 or 15-30) signal sequence #status predicted <SIG>
 F;31-491/Product: fibrinogen beta chain #status experimental <MAT>
 F;31-44/Product: fibrinopeptide B #status experimental <APT>
 F;45-491/Product: fibrin beta chain #status experimental <FCB>
 F;45-47/Region: polymerization site
 F;99-228/Domain: fibrinogen disulfide ring homology <FDR>
 F;238-487/Domain: fibrinogen beta/gamma homology <FBG>
 F;31/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimental
 F;44-45/Cleavage site: Arg-Gly (thrombin) #status experimental
 F;95/Disulfide bonds: interchain (to alpha-55) #status experimental
 F;106/Disulfide bonds: interchain (to alpha-68) #status experimental
 F;110/Disulfide bonds: interchain (to gamma-45) #status experimental
 F;223/Disulfide bonds: interchain (to alpha-184) #status experimental
 F;227/Disulfide bonds: interchain (to gamma-161) #status experimental
 F;231-316,241-270,424-437/Disulfide bonds: #status experimental
 F;394/Binding site: carbohydrate (Asn) (covalent) #status experimental

Alignment Scores:
 Pred. No.: 11.1 Length: 491
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.82% Indels: 0
 DB: 1 Gaps: 0

US-09-270-437D-6 (1-3412) x FGHUB (1-491)

QY 3171 CTTCTCTCGTGTGTTTCTCGTTAAA 3145
 |||||
 Db 21 LeuLeuLeuCysValPheLeuValLys 29

RESULT 8
 A29635
 N;Alternate names: human (fragment)
 C;Species: Homo sapiens (man)
 C;Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 05-Nov-1999
 C;Accession: A29635; G44256
 R;Kadonaga, J.T.; Garner, K.R.; Masiarz, F.R.; Tjian, R.
 Cell 51, 1079-1090, 1987
 A;Title: Isolation of cDNA encoding transcription factor Sp1 and functional analysis of
 A;Reference number: A29635; MUID:88080466; PMID:3319186
 A;Accession: A29635
 A;Molecule type: mRNA
 A;Residues: 1-696 <KAD>
 R;Ragoussis, J.; Senger, G.; Mockridge, I.; Sansseau, P.; Ruddy, S.; Dudley, K.; Sheer, D
 Genomics 14, 673-679, 1992
 A;Title: A testis-expressed Zn finger gene (ZNF76) in human 6p21.3 centromeric to the MH
 A;Reference number: A44256; MUID:93052398; PMID:1427894
 A;Accession: G44256
 A;Status: preliminary; not compared with conceptual translation
 A;Molecule type: nucleic acid
 A;Residues: 569-598 <RAG>
 A;Experimental source: T-cell line CEM
 A;Note: sequence extracted from NCBI backbone (NCBIP:125980)
 C;Genetics:
 A;Gene: GDB:SP1
 A;Cross-references: GDB:127453; OMIM:189906
 A;Map position: 19q13.1-19q13.3
 C;Keywords: DNA binding; transcription regulation; zinc finger

Alignment Scores:
 Pred. No.: 10.6 Length: 696
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.82% Indels: 0
 DB: 2 Gaps: 0

US-09-270-437D-6 (1-3412) x A29635 (1-696)

QY 893 ATCTCTTCGGTAGTTTGTCATCT 867
 |||||
 Db 216 IleSerSerAlaSerIleuValSerSer 224

RESULT 9
 JS0747
 regulatory protein Sp1 - rat
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 01-Dec-2000
 C;Accession: JS0747; S25287
 R;Imataka, H.; Sogawa, K.; Yasumoto, K.; Kikuchi, Y.; Sasano, K.; Kobayashi, A.; Hayami
 submitted to JIPID, September 1992
 A;Reference number: JS0747
 A;Accession: JS0747
 A;Molecule type: mRNA
 A;Residues: 1-788 <IMA>
 R;Imataka, H.; Sogawa, K.; Yasumoto, K.; Kikuchi, Y.; Sasano, K.; Kobayashi, A.; Hayami
 EMBO J. 11, 3663-3671, 1992
 A;Title: Two regulatory proteins that bind to the basic transcription element (BTE), a
 A;Reference number: S25287; MUID:93010958; PMID:1356762
 A;Accession: S25287
 A;Molecule type: mRNA
 A;Residues: 1-122,'L',124-311,'A',313-788 <IM2>
 C;Keywords: DNA binding; transcription regulation

Alignment Scores:
 Pred. No.: 10.4 Length: 788
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.82% Indels: 0
 DB: 2 Gaps: 0

US-09-270-437D-6 (1-3412) x JS0747 (1-788)

QY 893 ATCTCTTCGGTAGTTTGTCATCT 867
 |||||
 Db 308 IleSerSerAlaSerIleuValSerSer 316

RESULT 10
 T49915
 pre-mRNA splicing factor ATP-dependent RNA helicase-like protein - Arabidopsis thaliana
 N;Alternate names: protein T24H18.180
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
 C;Accession: T49915
 R;Bevan, M.; Robben, J.; Grymonprez, B.; Volckaert, G.; Bancroft, I.; Mewes, H.W.; Rueda
 submitted to the Protein Sequence Database, April 2000
 A;Reference number: Z25024
 A;Accession: T49915
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-1226 <BEV>
 A;Cross-references: EMBL:AL353013; GSPDB:GN00063; ATSP:T24H18.180
 C;Genetics:
 A;Map position: 5
 A;Introns: 2/3; 40/1; 111/1; 123/3; 183/3; 231/1; 279/3; 313/2; 349/3; 409/1; 485/3; 501/1
 A;Keywords: DNA binding; transcription regulation; zinc finger

Alignment Scores:
 Pred. No.: 9.8 Length: 1226
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.82% Indels: 0
 DB: 2 Gaps: 0

US-09-270-437D-6 (1-3412) x T49915 (1-1226)

Qy 499 GATGCCCGCTTAGCTTCTCCATGGCTA 473
|||||
Db 216 AspAlaArgLeuAlaSerProTrpLeu 224

RESULT 11
S52099
Phospholipase C beta 3 - human
N;Alternate names: phospholipase c beta-3, phosphoinositide-specific
C;Species: Homo sapiens (man)
C;Date: 15-Jul-1995 #sequence revision 21-Jul-1995 #text_change 17-Mar-1999
A;Accession: S52099; A56833; A56854
R;Lagercrantz, J.; Carson, E.; Phelan, C.; Grimmond, S.; Rosen, A.; Dare, E.; Nordenskjöld
submitted to the EMBL Data Library, September 1994
A;Description: Genomic organization and complete cDNA sequence of the human phosphoinosi
A;Reference number: S52099
A;Accession: S52099
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1234 <LAG>
A;Cross-references: EMBL:Z37573
R;Lagercrantz, J.; Carson, E.; Phelan, C.; Grimmond, S.; Rosen, A.; Dare, E.; Nordenskjöld
Genomics 26, 467-472, 1995
A;Title: Genomic organization and complete cDNA sequence of the human phosphoinositide-s
A;Reference number: A56833; MUID:95331781; PMID:7607669
A;Accession: A56833
A;Molecule type: DNA
A;Residues: 1-200 <LA2>
A;Cross-references: GB:Z37544
R;Sinke, R.J.; Geurts van Kessel, A.
Genomics 25, 568-569, 1995
A;Title: Localization of the human phosphatidylinositol-specific phospholipase C beta-3
A;Reference number: A56854; MUID:95309927; PMID:7789993
A;Accession: A56854
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-193 <SIN>
C;Genetics:
A;Gene: GDB:PLCB3
A;Cross-references: GDB:386061; OMIM:600230
A;Map position: 11q13-11q13
A;Introns: 33/3; 59/3; 82/3; 129/3; 156/2; 174/2; 199/3; 233/2; 288/3; 338/1; 418/2; 446
1119/2; 1138 /3; 1167/3
C;Superfamily: 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase I; 1-phosphatid
odiesterase domain Y homology
F;319-468/Domain: 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase domain X hom
F;589-709/Domain: 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase domain Y hom

Alignment Scores:
Pred. No.: 9.79 Length: 1234
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.82% Indels: 0
DB: 2 Gaps: 0

US-09-270-437D-6 (1-3412) x S52099 (1-1234)

Qy 217 AGTTCTGTCGGGTAGTCACGAAGG 191
|||||
Db 1099 SerSerGlyArgGlySerProArgArg 1107

RESULT 12
I73957
kinase-related protein c-ros-1 precursor - rat
N;Contains: protein-tyrosine kinase (EC 2.7.1.112) ros-1
C;Species: Rattus norvegicus (Norway rat)
C;Date: 02-Aug-1996 #sequence revision 02-Aug-1996 #text_change 02-Jun-2000
C;Accession: I73957; I56752; I73956
R;Matsushima, H.; Shibuya, M.
J. Virol. 64, 2117-2125, 1990
A;Title: Tissue-specific expression of rat c-ros-1 gene and partial structural similarit
A;Reference number: I56752; MUID:90219211; PMID:2139140

A;Accession: I73957
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: mRNA
A;Residues: 1-2338 <RES>
A;Cross-references: GB:M35106; NID:g203599; PIDN:AAA40968.1; PID:g203600
A;Accession: I56752
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: mRNA
A;Residues: 1-430,452-2338 <RE2>
A;Cross-references: GB:M35104; NID:g203595; PIDN:AAA40966.1; PID:g203596
A;Accession: I73956
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: mRNA
A;Residues: 1-430,452-1872,'AC',1875 <RE3>
A;Cross-references: GB:M35105; NID:g203597; PIDN:AAA40967.1; PID:g203598
C;Superfamily: kinase-related protein ros; LDL receptor YWTD-containing repeat homology
C;Keywords: alternative splicing; ATP; autophosphorylation; glycoprotein; kinase-relate
ific protein kinase
F;753-793/Domain: LDL receptor YWTD-containing repeat homology <YW3>
F;1335-2214/Domain: protein kinase homology <KIN>
F;1943-1951/Region: protein kinase ATP-binding motif

Alignment Scores:
Pred. No.: 8.98 Length: 2338
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 2 Gaps: 0

US-09-270-437D-6 (1-3412) x I73957 (1-2338)

Qy 569 AGCGAGCCGCGTGGGACACCTCTT 595
|||||
Db 656 SerGluProSerValGlyThrThrLeu 664

RESULT 13
TVHURS
kinase-related protein ros-1 precursor - human
N;Alternate names: protein-tyrosine kinase mcf3 (activated ros-1)
N;Contains: protein-tyrosine kinase (EC 2.7.1.112) ros-1
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1988 #sequence revision 07-Oct-1994 #text_change 11-Jun-1999
A;Accession: A35512; A24223; A24421; A33081
R;Birchmeier, C.; O'Neill, K.; Riggs, M.; Wigler, M.
Proc. Natl. Acad. Sci. U.S.A. 87, 4799-4803, 1990
A;Title: Characterization of ROS1 cDNA from a human glioblastoma cell line.
A;Reference number: A35512; MUID:90280463; PMID:2352949
A;Accession: A35512
A;Molecule type: mRNA
A;Residues: 1-2212,'N',2214-2227,'QC',2229-2347 <BIR>
A;Cross-references: GB:M34353
A;Experimental source: glioblastoma cell line SW-1088
R;Matsushima, H.; Wang, L.H.; Shibuya, M.
Mol. Cell. Biol. 6, 3000-3004, 1986
A;Title: Human c-ros-1 gene homologous to the v-ros sequence of UR2 sarcoma virus encod
A;Reference number: A25223; MUID:87064611; PMID:3023956
A;Accession: A25223
A;Molecule type: DNA
A;Residues: 1790-2245,'KFDGSEPSFRCTVN' <MA2>
A;Cross-references: GB:M13368
A;Experimental source: placenta
A;Note: the differences after residue 2245 result from the authors' misinterpretation of
R;Birchmeier, C.; Birnbaum, D.; Waitches, G.; Fasano, O.; Wigler, M.
Mol. Cell. Biol. 6, 3109-3116, 1986
A;Title: Characterization of an activated human ros gene.
A;Reference number: A24421; MUID:87064625; PMID:3785223
A;Accession: A24421
A;Molecule type: mRNA
A;Residues: 1854-2261,'A',2263-2347 <BI2>
A;Cross-references: GB:M13880; NID:g337482; PIDN:AAA36580.1; PID:g337483
A;Experimental source: tumor cells
A;Note: the mcf3 oncogene was formed by DNA rearrangement involving fusion of at least t

C:Genetics:
A:Gene: GDB:ROS1
A:Cross-references: GDB:120351; OMIM:165020
A:Map position: 6q22-6q22
C:Superfamily: Kinase-related protein tos; LDL receptor YWTD-containing repeat homology;
C:Species: ATP; autophosphorylation; glycoprotein; kinase-related transforming protein;
C:Keywords: ATP; autophosphorylation; signal sequence #status predicted <SIG>
F:1-36/Domain: kinase-related protein ROS1 #status predicted <MAT>
F:37-2347/Product: kinase-related protein ROS1 #status predicted <MAT>
F:37-1859/Domain: extracellular #status predicted <EXT>
F:335-378/Domain: LDL receptor YWTD-containing repeat homology <YW1>
F:466-503/Domain: LDL receptor YWTD-containing repeat homology <YWA>
F:715-757/Domain: LDL receptor YWTD-containing repeat homology <YW2>
F:758-798/Domain: LDL receptor YWTD-containing repeat homology <YW3>
F:799-838/Domain: LDL receptor YWTD-containing repeat homology <YW4>
F:843-888/Domain: LDL receptor YWTD-containing repeat homology <YW5>
F:893-933/Domain: LDL receptor YWTD-containing repeat homology <YW6>
F:1530-1574/Domain: LDL receptor YWTD-containing repeat homology <YW7>
F:1860-1883/Domain: transmembrane #status predicted <TMN>
F:1884-2347/Domain: intracellular #status predicted <INT>
F:1943-2222/Domain: protein kinase homology <KIN>
F:1951-1959/Region: protein kinase ATP-binding motif
F:52,114,123,324,352,471,607,628,706,714,732,939,961,1015,1087,1090,1211,1272,1330,1458,
F:1980/Active site: Lys #status predicted
F:2110,2114,2115/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #stat

Alignment Scores:
Pred. No.: 8.98 Length: 2347
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 1 Gaps: 0

US-09-270-437D-6 (1-3412) x TVHURS (1-2347)

QY 569 ACCGAGCCAGCGTGGGACCACTCTT 595
Db 661 SerGluProSerValGlyThrThrLeu 669
|||||

RESULT 14
T12121
NADH dehydrogenase 4 - Atlantic horseshoe crab mitochondrion (fragment)
C:Species: mitochondrion Limulus polyphemus (Atlantic horseshoe crab)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 11-May-2000
C:Accession: T12121
R:Staton, J.L.; Daehler, L.L.; Brown, W.M.
Mol. Biol. Evol. 14, 867-874, 1997
A:Title: Mitochondrial gene arrangement of the horseshoe crab Limulus polyphemus L.: Con
A:Reference number: Z17427; MUID:97398711; PMID:9254925
A:Accession: T12121
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-63 <STA>
A:Cross-references: EMBL:AF002647; NID:g2316048; PID:g2316055; PIDN:AAC47689.1
C:Genetics:
A:Genome: mitochondrion
A:Note: ND4
C:Keywords: mitochondrion

Alignment Scores:
Pred. No.: 144 Length: 63
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.73% Indels: 0
DB: 2 Gaps: 0

US-09-270-437D-6 (1-3412) x T12121 (1-63)

QY 1833 TTTCGTCTCTGCTGCTTCACCTG 1810
Db 39 PheLeuLeuLeuLeuHisLeu 46
|||||

RESULT 15

F72580
hypothetical protein APE1925 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C:Accession: F72580
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Taka
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Cross-references: DDBJ:AP000062; NID:g5105244; PIDN:BAA80931.1; PID:d1044717; PID:g51
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE1925

Alignment Scores:
Pred. No.: 132 Length: 121
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.73% Indels: 0
DB: 2 Gaps: 0

US-09-270-437D-6 (1-3412) x F72580 (1-121)

QY 2058 CTTCCCGCCCTCTCGGCCCT 2035
Db 95 LeuProArgProSerSerAlaPro 102
|||||

Search completed: July 16, 2004, 11:32:56
Job time : 82.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: July 16, 2004, 10:45:36 ; Search time 32 Seconds
(without alignments)
11103.961 Million cell updates/sec

Title: US-09-270-437D-6

Perfect score: 1120

Sequence: 1 ggcaggaggaggcgagga.....aaccttgaaaattttattt 3412

Scoring table: OLIGO

Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 141681 seqs, 52070155 residues

Word size: 1

Total number of hits satisfying chosen parameters: 283186

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODE=frame*_n2p.model -DRV=xlp
-Q=/cgn2_1/USPTO.spool_p/US09270437/rumat_16072004_113127_13943/app.query.fasta_1.3591
-DB=SwissProt 42 -QWMT=fastan -SUFFIX=Oligo.rsp -MINMATCH=0.1 -LOOPTCL=0
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptco -NORM=ext -HEA=SIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09270437.acgn_1_1_46_rumat_16072004_113127_13943 -NCPUS=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGIOG
-DRV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	0.8	398	1 KR2_PRVN3	P30662 pseudorabie
2	9	0.8	491	1 FIBR_HUMAN	P02675 homo sapien
3	9	0.8	781	1 SPI_MOUSE	O89090 mus musculu
4	9	0.8	785	1 SPI_HUMAN	P08047 homo sapien
5	9	0.8	788	1 SPI_RAT	Q01714 rattus norv
6	9	0.8	2347	1 KROS_HUMAN	P08922 homo sapien
7	8	0.7	111	1 FTSE_RALSO	O8y0b4 raltonia s
8	8	0.7	137	1 N551_SOYBN	Q05544 glycine max
9	8	0.7	150	1 SHIB_PIG	P79399 sus scrofa
10	8	0.7	161	1 SHIB_CANFA	P79250 canis famil
11	8	0.7	171	1 Y175_STAAM	O53719 staphylococ
12	8	0.7	202	1 YS2_NEUCR	O43105 neurospora
13	8	0.7	208	1 YS2_CAEEL	Q10021 caenorhabd
14	8	0.7	224	1 RS2_METMA	O8pw41 methanosarc
15	8	0.7	225	1 RS2_METAC	O8tt39 methanosarc
16	8	0.7	244	1 CSMI_CHLTE	O68988 chlorobium
17	8	0.7	285	1 Y891_MYCTU	Q10551 mycobacteri
18	8	0.7	285	1 Y915_MYCBO	P59970 mycobacteri

C 19	8	0.7	303	1 GDFF_MOUSE	Q92017 mus musculu
C 20	8	0.7	303	1 GDFF_RAT	Q92016 rattus norv
C 21	8	0.7	311	1 RBFI_MOUSE	P22560 mus musculu
C 22	8	0.7	316	1 TAL_HELPJ	O92jc5 helicobacte
C 23	8	0.7	326	1 YJHS_ECOLI	P39370 escherichia
C 24	8	0.7	379	1 CYB_PENFU	Q9blx4 pentagalus
C 25	8	0.7	381	1 CYB_NOTTY	O03478 notoryctes
C 26	8	0.7	386	1 SHIB_CRIGR	P46636 cricetus
C 27	8	0.7	386	1 SHIB_MOUSE	P28334 mus musculu
C 28	8	0.7	386	1 SHIB_RAT	P28564 rattus norv
C 29	8	0.7	386	1 SHIB_SPAEH	P56496 spalax leuc
C 30	8	0.7	388	1 SHIB_DIDMA	P35404 didelphis m
C 31	8	0.7	389	1 SHIB_CAVPO	O08892 cavia porce
C 32	8	0.7	390	1 SHIB_HUMAN	P28222 homo sapien
C 33	8	0.7	390	1 SHIB_PANTR	P60020 pan troglod
C 34	8	0.7	390	1 SHIB_RABIT	P49144 coryctolagus
C 35	8	0.7	395	1 UMPI_ARATH	O9lka5 arabidopsis
C 36	8	0.7	407	1 IE68_HSVSA	Q01042 herpesvirus
C 37	8	0.7	411	1 IHH_MOUSE	P97812 mus musculu
C 38	8	0.7	414	1 SX17_HUMAN	Q9hg12 homo sapien
C 39	8	0.7	419	1 SX17_MOUSE	O61473 mus musculu
C 40	8	0.7	428	1 FXB2_MOUSE	O64733 mus musculu
C 41	8	0.7	428	1 HEMY_HAEIN	P44772 haemophilus
C 42	8	0.7	480	1 PRCP_HUMAN	P10619 homo sapien
C 43	8	0.7	498	1 VE2_HPVO8	P06422 human papil
C 44	8	0.7	499	1 PITB_ECOLI	P43676 escherichia
C 45	8	0.7	500	1 CP46_HUMAN	Q9y6a2 homo sapien

ALIGNMENTS

RESULT 1

KR2_PRVN3	KR2_PRVN3	STANDARD;	PRT;	398 AA.
AC	P30662;			
DT	01-APR-1993 (Rel. 25, Created)			
DT	01-APR-1993 (Rel. 25, Last sequence update)			
DT	01-OCT-1996 (Rel. 34, Last annotation update)			
DE	Serine/threonine-protein kinase 2 (EC 2.7.1.1.-)			
GN	UL13 OR ULPK.			
OS	Pseudorabies virus (strain NIA-3) (PRV).			
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;			
OC	Alphaherpesvirinae; Varicelloviruses.			
OX	NCBI_Taxid=10349;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92365105; PubMed=1323689;			
RA	de Wind N., Domen J., Berns A.;			
RT	"Herpesviruses encode an unusual protein-serine/threonine kinase			
RT	which is nonessential for growth in cultured cells.";			
RL	J. Virol. 66:5200-5209(1992).			
CC	-!- FUNCTION: PROTEIN-SERINE/THREONINE KINASE NONESSENTIAL FOR GROWTH			
CC	IN CULTURED CELLS.			
CC	-!- PTM: Autophosphorylated (possible).			
CC				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC				
DR	EMBL; M94870; AAA47481.1; -			
DR	PIR; B42744; WZBEN3.			
DR	InterPro; IPR000719; Prot kinase.			
DR	InterPro; IPR008271; Ser_thr_pkin_AS.			
DR	PROSITE; PS00107; PROTEIN KINASE ATP; FALSE_NEG.			
DR	PROSITE; PS00108; PROTEIN KINASE ST; 1.			
DR	PROSITE; PS50011; PROTEIN KINASE_DOM; 1.			
KW	Transferase; Serine/threonine-protein kinase; ATP-binding;			
KW	Phosphorylation; Manganese.			
FT	DOMAIN 4 9 POLY-GLY.			

FT DOMAIN 80 398 PROTEIN KINASE.
 FT NP BIND 86 94 ATP (BY SIMILARITY).
 FT BINDING 103 103 ATP (BY SIMILARITY).
 FT ACT SITE 194 194 BY SIMILARITY.
 SQ SEQUENCE 398 AA; 41416 MW; 41416 MW; D5C69AD75E42309B CRC64;

Alignment Scores:
 Pred. No.: 4.87 Length: 398
 Score: 3.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.80% Indels: 0
 DB: 1 Gaps: 0

US-09-270-437D-6 (1-3412) x KR2_PRVN3 (1-398)

Qy 2 GCAGCGGAGGAGCGAGGAGCGCGGG 28

Db 119 AlaxlaGluGluAlaArgSerAlaGly 127

RESULT 2

FIBB HUMAN

ID FIBB_HUMAN STANDARD; PRT; 491 AA.

AC P02675;

DT 21-JUL-1986 (Rel. 01, Created)

DT 01-JUL-1993 (Rel. 26, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Fibrinogen beta chain precursor [Contains: Fibrinopeptide B].

GN FGB.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RN SEQUENCE FROM N.A.

RX MEDLINE=91344740; PubMed=2102623;

RA Chung D.W., Harris J.E., Davie E.W.;

RT "Nucleotide sequences of the three genes coding for human

fibrinogen";

RL Adv. Exp. Med. Biol. 281:39-48(1990).

RN [2]

RN SEQUENCE FROM N.A.

RX MEDLINE=83283433; PubMed=6688356;

RA Chung D.W., Que B.G., Rixon M.W., Mace M. Jr., Davie E.W.;

RT "Characterization of complementary deoxyribonucleic acid and genomic

deoxyribonucleic acid for the beta chain of human fibrinogen.";

RL Biochemistry 22:3244-3250(1983).

RN [3]

RN SEQUENCE FROM N.A.

RA Chung D.W., Harris J.E., Davie E.W.;

RT "Nucleotide sequences of the three genes coding for human

fibrinogen.";

RL (in) Liu C.Y., Chien S. (eds.);

PLenum Press, New York (1991).

RN [4]

RN SEQUENCE FROM N.A., AND VARIANTS SER-100; HIS-170; LEU-265 AND

LYS-478.

RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,

RA Nickerson D.A.;

RL Submitted (JUN-2001) to the EMBL/GenBank/DBPJ databases.

RN [5]

RN SEQUENCE OF 31-491, AND CARBOHYDRATE-LINKAGE SITE.

RA Henschen A., Lottspeich F., Southan C., Topfer-Petersen E.;

RT "Human fibrinogen: sequence, sulfur bridges, glycosylation and some

structural variants.";

RL (in) Peeters H. (eds.);

PL Protides of the biological fluids, Proc. 28th colloquium, pp.51-56,

RL Pergamon Press, Oxford (1980).

RN [6]

RN SEQUENCE OF 31-491.

RX MEDLINE=79124640; PubMed=420779;

RA Watt K.W.K., Takagi T., Doolittle R.F.;

RT "Amino acid sequence of the beta chain of human fibrinogen.";
 RL Biochemistry 18:68-76(1979).
 RN [7]
 RP SEQUENCE OF 31-148, AND DISULFIDE BONDS.
 RX MEDLINE=76225080; PubMed=936108;
 RA Blomback B., Hessel B., Hogg D.;

RT "Disulfide bridges in NH2-terminal part of human fibrinogen.";

RL Thromb. Res. 8:639-658(1976).

RN [8]

RN SEQUENCE OF 1-38 FROM N.A.

RX MEDLINE=87146483; PubMed=3029722;

RA Huber P., Dalmon J., Courtois G., Laurent M., Assouline Z.,

RA Marguerie G.;

RT "Characterization of the 5'-flanking region for the human fibrinogen

beta gene.";

RL Nucleic Acids Res. 15:1615-1625(1987).

RN [9]

RN SEQUENCE OF 31-44.

RA Blomback B., Blomback M., Grondahl N.J., Guthrie C., Hinton M.;

RT "Studies on fibrinopeptides from primates.";

RL Acta Chem. Scand. 19:1788-1789(1965).

RN [10]

RP REVIEW, AND DISULFIDE BONDS.

RX MEDLINE=83254370; PubMed=6575689;

RA Henschen A., Lottspeich F., Kehl M., Southan C.;

RT "Covalent structure of fibrinogen.";

RL Ann. N.Y. Acad. Sci. 408:28-43(1983).

RN [11]

RP DISULFIDE BONDS.

RX MEDLINE=77245999; PubMed=891553;

RA Gaardlund B., Hessel B., Marguerie G., Murano G., Blomback B.;

RT "Primary structure of human fibrinogen. Characterization of

disulfide-containing cyanogen-bromide fragments.";

RL Eur. J. Biochem. 77:595-610(1977).

RN [12]

RP DISULFIDE BONDS.

RA Doolittle R.F., Takagi T., Watt K.W.K., Bouma H. III, Cottrell B.A.,

RA Cassman K.G., Goldbaum D.M., Doolittle L.R., Friezner S.J.;

RT "The structures of fibrinogen and fibrin.";

RL (in) Magnusson S., Ottesen M., Foltmann B., Dano K.,

RL Neurath H. (eds.);

RL Regulatory proteolytic enzymes and their inhibitors, pp.163-172,

RL Pergamon Press, New York (1978).

RN [13]

RP REVIEW, EM STRUCTURE, POLYMERIZATION, AND LIGANDS.

RX MEDLINE=84305751; PubMed=6383194;

RA Doolittle R.F.;

RT "Fibrinogen and fibrin.";

RL Annu. Rev. Biochem. 53:195-229(1984).

RN [14]

RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF 164-491.

RX MEDLINE=97472408; PubMed=933233;

RA Spraggon G., Everse S.J., Doolittle R.F.;

RT "Crystal structures of fragment D from human fibrinogen and its

crosslinked counterpart from fibrin.";

RL Nature 389:455-462(1997).

RN [15]

RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 164-491.

RX MEDLINE=98292395; PubMed=9628725;

RA Everse S.J., Spraggon G., Veerapandian L., Riley M., Doolittle R.F.;

RT "Crystal structure of fragment Double-D from human fibrin with two

different bound ligands.";

RL Biochemistry 37:8637-8642(1998).

RN [16]

RP X-RAY CRYSTALLOGRAPHY.

RX MEDLINE=99175089; PubMed=10074346;

RA Everse S.J., Spraggon G., Veerapandian L., Doolittle R.F.;

RT "Conformational changes in fragments D and double-D from human

fibrin(ogen) upon binding the peptide ligand Gly-His-Arg-Pro-amide.";

RL Biochemistry 38:2941-2946(1999).

RN [17]

RP INTERACTION WITH FBLN1.

RX MEDLINE=95370284; PubMed=7642629;

RA Tran H., Tanaka A., Litvinovich S.V., Medved L.V., Haudenschild C.C.,
RA Atgraves W.S.;
RT "The interaction of fibulin-1 with fibrinogen. A potential role in
RT hemostasis and thrombosis.";
RL J. Biol. Chem. 270:19458-19464(1995).
RN [18]
RP VARIANT BALTIMORE-2 LYS-478.
RX MEDLINE=89058942; PubMed=3194892;
RA Schmelzer C.H., Ebert R.F., Bell W.R.;
RT "A polymorphism at B beta 448 of fibrinogen identified during
RT structural studies of fibrinogen Baltimore II.";
RL Thromb. Res. 52:173-177(1988).
RN [19]
RP VARIANT ISE ARG-45.
RX MEDLINE=91208409; PubMed=2018836;
RA Yoshida N., Wada H., Morita K., Hirata H., Matsuda M., Yamazumi K.,
RA Asakura S., Shirakawa S.;
RT "A new congenital abnormal fibrinogen Ise characterized by the
RT replacement of B beta glycine-15 by cysteine.";
RL Blood 77:1958-1963(1991).
RN [20]
RP VARIANT NAPLES THR-98.
RX MEDLINE=92340664; PubMed=1634610;
RA Koopman J., Haverkate F., Lord S.T., Grimbergen J., Mannucci P.M.;
RT "Molecular basis of fibrinogen Naples associated with defective
RT thrombin binding and thrombophilia. Homozygous substitution of B beta
RT 68 Ala-->Thr.";
RL J. Clin. Invest. 90:238-244(1992).
RN [21]
RP VARIANTS IJMUUDEN CYS-44 AND NIJMEGEN CYS-74.
RX MEDLINE=92228809; PubMed=1565641;
RA Koopman J., Haverkate F., Grimbergen J., Engesser L., Novakova I.,
RA Kerst A.F.J.A., Lord S.T.;
RT "Abnormal fibrinogens IJmuuden (B beta Arg14-->Cys) and Nijmegen (B
RT beta Arg44-->Cys) form disulfide-linked fibrinogen-albumin
RT complexes.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:3478-3482(1992).
RN [22]
RP VARIANT NEW YORK-1 39-GLY--LEU-102 DEL.
RX MEDLINE=85157605; PubMed=3156856;
RA Liu C.Y., Koehn J.A., Morgan F.J.;
RT "Characterization of fibrinogen New York 1. A dysfunctional
RT fibrinogen with a deletion of B beta(9-72) corresponding exactly to
RT exon 2 of the gene.";
RL J. Biol. Chem. 260:4390-4396(1985).
RN [23]
RP VARIANTS GLU-2; LEU-265 AND LYS-478.
RX MEDLINE=99318093; PubMed=10391209;
RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
RA Lander E.S.;
RT "Characterization of single-nucleotide polymorphisms in coding regions
RT of human genes.";
RL Nat. Genet. 22:231-238(1999).
RN [24]
RP ERRATUM.
RA Cargill M., Altschuler D., Ireland J., Sklar P., Ardlie K., Patil N.,
RA Shaw N., Lane C.R., Lim E.P., Kalyanaraman N., Nemesh J., Ziaugra L.,
RA Friedland L., Rolfe A., Warrington J., Lipshutz R., Daley G.Q.,
RA Lander E.S.;
RL Nat. Genet. 23:373-373(1999).
RN [25]
RP VARIANTS CONGENITAL AFIBRINOGENEMIA ARG-383 AND ASP-430.
RX MEDLINE=20129589; PubMed=10666208;
RA Duga S., Asselta R., Santagostino E., Zeinali S., Simonc T.,
RA Malcovati M., Mannucci P.M., Turchini M.L.;
RT "Missense mutations in the human beta fibrinogen gene cause
RT congenital afibrinogenemia by impairing fibrinogen secretion.";
RL Blood 95:1336-1341(2000).
RN [26]
RP VARIANT CONGENITAL AFIBRINOGENEMIA CYS-196.
RX MEDLINE=21361164; PubMed=11468164;

RA Lounes K.C., Lefkowitz J.B., Henschen-Edman A.H., Coates A.I.,
RA Hantgan R.R., Lord S.T.;
RT "The impaired polymerization of fibrinogen Longmont
RT Alignment Scores:
Pred. No.: 4.79 Length: 491
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.82% Indels: 0
DB: 1 Gaps: 0
US-09-270-437D-6 (1-3412) x FIBB_HUMAN (1-491)
QY 3171 CTTCTCTGTGTGTTTTCTCGTTAAA 3145
Db 21 LeuLeuLeuCysValPheLeuValLys 29
RESULT 3
SPI_MOUSE
ID SPI_MOUSE STANDARD; PRT; 781 AA.
AC O89090; Q62251; Q64167;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Transcription factor Sp1.
GN SP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Neuroblastoma;
RX MEDLINE=98290554; PubMed=9628590;
RA Vajina S., Lee S.H., Minowa T., Mouradian M.M.;
RT "Sp family transcription factors regulate expression of rat D2
RT dopamine receptor gene.";
RL DNA Cell Biol. 17:471-479(1998).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=96016118; PubMed=7568082;
RA Persengiev S.P., Saffer J.D., Kilpatrick D.L.;
RT "An alternatively spliced form of the transcription factor Sp1
RT containing only a single glutamine-rich transactivation domain.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:9107-9111(1995).
RN [3]
RP SEQUENCE OF 681-781 FROM N.A.
RX MEDLINE=92338398; PubMed=1633330;
RA Chestier A., Charney P.;
RT "Difference in the genomic organizations of the related transcription
RT factors Sp1 and Krox-20; possible evolutionary significance.";
RL DNA Seq. 2:325-327(1992).
CC -!- FUNCTION: BINDS TO GC BOX PROMOTERS ELEMENTS AND SELECTIVELY
CC ACTIVATES MRNA SYNTHESIS FROM GENES THAT CONTAIN FUNCTIONAL
CC RECOGNITION SITES.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=O89090-1; Sequence=Displayed;
CC Name=2; Synonyms=Sp1-S;
CC IsoId=O89090-2; Sequence=VSP 007376;
CC -!- PTM: O-glycosylated; contains N-acetylglucosamine side chains (By
CC similarity).
CC -!- SIMILARITY: BELONGS TO THE SP1 FAMILY OF C2H2-TYPE ZINC-FINGER
CC PROTEINS.
CC -!- SIMILARITY: Contains 3 C2H2-type zinc fingers.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; AF062566; AAC16484.1; -.
 DR EMBL; S79832; AAB35321.1; -.
 DR EMBL; X60136; CAA42721.1; -.
 DR HSPF; P08047; 1SP1.
 DR MGO; MGI:98372; Spt.
 DR GO; GO:0005634; C:nucleus; IDA.
 DR GO; GO:0003700; F:transcription factor activity; IMP.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IMP.
 DR InterPro; IPR007087; Znf_C2H2.
 DR Pfam; PF00096; zf-C2H2_3.
 DR ProDom; PD000003; Znf_C2H2; 2.
 DR SMART; SM00335; Znf_C2H2; 3.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
 DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 3.
 KW Transcription regulation; Activator; Zinc-finger; Metal-Binding;
 KW DNA-binding; Nuclear protein; Repeat; Glycoprotein;
 KW Alternative splicing.
 FT ZN_FING 624 648 C2H2-TYPE 1.
 FT ZN_FING 654 678 C2H2-TYPE 2.
 FT ZN_FING 684 706 C2H2-TYPE 3.
 FT VARSPLIC 57 370 Missing (in isoform 2).
 FT /FTId=Vsp_007376.
 FT CONFLICT 459 459 V -> G (IN REF. 2).
 SQ SEQUENCE 781 AA; 80486 MW; 14CD12B8C58CF921 CRC64;

Alignment Scores:
 Pred. No.: 4,6 Length: 781
 Score: 9,00 Matches: 9
 Percent Similarity: 100,00% Conservative: 0
 Best Local Similarity: 100,00% Mismatches: 0
 Query Match: 0,82% Indels: 0
 DB: 1 Gaps: 0

US-09-270-437D-6 (1-3412) x SPL_MOUSE (1-781)
 QY 893 ATCTCTTCGGCTAGTTTGGTTCATCT 867
 DB 304 IleSerSerAlaSerLeuValSerSer 312

RESULT 4
 ID SPL_HUMAN STANDARD; PRT; 785 AA.
 AC P08047; Q9H305; Q9NR51; Q9NY21; Q9NYE7;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Transcription factor Spt.
 GN SPL OR TSFPL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE OF 4-785 FROM N.A.
 RC TISSUE=Cervical carcinoma;
 RA Haggart M.H., Ladurner A.G.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 1-558 FROM N.A.
 RX MEDLINE=20545561; PubMed=10973950;
 RA Takahara T., Kanazu S., Yanagisawa S., Akanuma H.;
 RT "Heterogeneous Spt mRNAs in human HepG2 cells include a product of
 RT homotypic trans-splicing."
 RL J. Biol. Chem. 275:38067-38072 (2000).
 RN [3]
 RP SEQUENCE OF 90-785 FROM N.A., AND SEQUENCE OF 359-375 AND 670-675.
 RX MEDLINE=8980466; PubMed=319186;
 RA Kadonaga J.T., Carner K.R., Masiaz F.R., Tjian R.;
 RT "Isolation of cDNA encoding transcription factor Spt and functional

analysis of the DNA binding domain.";
 Cell 51:1079-1090(1987).
 [4]
 RP SEQUENCE OF 1-109 FROM N.A.
 RA Nicolas M., Noe V., Ciudad C.J.;
 RT "Expression of transcription factor Spt mRNA in mammalian cells."
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 [5]
 RP SEQUENCE OF 1-98 FROM N.A.
 RA Handschug K., Huebner A.;
 RT "Sequencing of the 5' end of human transcription factor Spt mRNA.";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 [6]
 RP O-GLYCOSYLATION.
 RX MEDLINE=89003041; PubMed=3139301;
 RA Jackson S.P., Tjian R.;
 RT "O-glycosylation of eukaryotic transcription factors: implications
 RT for mechanisms of transcriptional regulation.";
 RL Cell 55:125-133(1988).
 [7]
 RP STRUCTURE BY NMR OF 654-684 AND 684-712.
 RX MEDLINE=97218212; PubMed=9065444;
 RA Narayan V.A., Kriwacki R.W., Caradonna J.P.;
 RT "Structures of zinc finger domains from transcription factor Spt.
 RT Insights into sequence-specific protein-DNA recognition.";
 RL J. Biol. Chem. 272:7801-7809(1997).
 [8]
 RP IDENTIFICATION OF SEROTONIN 1A RECEPTOR PROMOTER BINDING SITES.
 RX MEDLINE=96224025; PubMed=8626793;
 RA Parks C.L., Shenk T.;
 RT "The serotonin 1a receptor gene contains a TATA-less promoter that
 RT responds to MAZ and Spt.";
 RL J. Biol. Chem. 271:4417-4430(1996).
 CC -!- FUNCTION: BINDS TO GC BOX PROMOTERS ELEMENTS AND SELECTIVELY
 CC ACTIVATES MRNA SYNTHESIS FROM GENES THAT CONTAIN FUNCTIONAL
 CC RECOGNITION SITES. CAN INTERACT WITH G/C-RICH MOTIFS FROM
 CC SEROTONIN RECEPTOR PROMOTER.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- PTM: O-glycosylated; contains N-acetylglucosamine side chains.
 CC -!- SIMILARITY: BELONGS TO THE SPT FAMILY OF C2H2-TYPE ZINC-FINGER
 CC PROTEINS.
 CC -!- SIMILARITY: Contains 3 C2H2-type zinc fingers.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF252284; AAF67726.1; -.
 DR EMBL; AB039286; BAB13476.1; -.
 DR EMBL; J03133; AAA61154.1; -.
 DR EMBL; AF255682; AAF78781.1; -.
 DR EMBL; AJ272134; CAB75345.1; -.
 DR PIR; A29635; A29635.
 DR PDB; 1SP1; 21-APR-97.
 DR PDB; 1SP2; 21-APR-97.
 DR TRANSFAC; T00759; -.
 DR GlycoSuiteDB; P08047; -.
 DR Genew; HGNC:11205; SPT.
 DR MIM; 189906; -.
 DR GO; GO:0005634; C:nucleus; NAS.
 DR GO; GO:0003702; F:RNA polymerase II transcription factor acti. . .; NAS.
 DR GO; GO:0016563; F:transcriptional activator activity; NAS.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
 DR InterPro; IPR007087; Znf_C2H2.
 DR Pfam; PF00096; zf-C2H2; 3.
 DR ProDom; PD000003; Znf_C2H2; 2.
 DR SMART; SM00335; Znf_C2H2; 3.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_2; 3.
 DR PROSITE; PS00157; ZINC_FINGER_C2H2_3; 3.

KW Transcription regulation; Activator; Zinc-finger; Metal-binding;
KW DNA-binding; Nuclear protein; Repeat; Glycoprotein; 3D-structure.
FT ZN FING 626 650 C2H2-TYPE 1.
DR ZN FING 656 680 C2H2-TYPE 2.
FT ZN FING 686 708 C2H2-TYPE 3.
FT ZN FING 686 708 C2H2-TYPE 3.
FT CONFLICT 366 366 D -> G (IN REF. 3; AA SEQUENCE).
FT CONFLICT 670 670 S -> F (IN REF. 3; AA SEQUENCE).
FT STRAND 657 657
FT TURN 661 662
FT STRAND 666 666
FT HELIX 670 677
FT TURN 678 680
SQ SEQUENCE 785 AA; 80693 MW; 43893DBP6518B9EA CRC64;

Alignment Scores:
Pred. No.: 4.6 Length: 785
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.82% Indels: 0
DB: 1 Gaps: 0

US-09-270-437D-6 (1-3412) x SP1_HUMAN (1-785)

QY 893 ATCTCTCGCTAGTTGGTCTCATCT 867
Db 305 IIESSerSAlaSerLeuValSerSer 313
|||||

RESULT 5

SP1_RAT
ID SP1_RAT STANDARD; PRT; 788 AA.
AC Q01714;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Transcription factor Sp1.
GN SP1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP TISSUE=Liver;
RC MEDLINE=93010958; PubMed=1356762;
RA Imataka H., Sogawa K., Yasumoto K., Kikuchi Y., Sasano K.,
RA Kobayashi A., Hayami M., Fujii-Kuriyama Y.;
RT "Two regulatory proteins that bind to the basic transcription element
(BTE), a GC box sequence in the promoter region of the rat P-4501A1
gene.";
RL ENBO J. 11:3663-3671(1992).
CC -!- FUNCTION: BINDS TO GC BOX PROMOTERS ELEMENTS AND SELECTIVELY
CC ACTIVATES MRNA SYNTHESIS FROM GENES THAT CONTAIN FUNCTIONAL
CC RECOGNITION SITES.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- PTM: O-glycosylated; contains N-acetylglucosamine side chains (By
CC similarity).
CC -!- SIMILARITY: BELONGS TO THE SP1 FAMILY OF C2H2-TYPE ZINC-FINGER
CC PROTEINS.
CC -!- SIMILARITY: Contains 3 C2H2-type zinc fingers.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D12768; BAA02235.1; -.
CC PIR; JS0747; JS0747.
CC HSSP; P08047; 1SP1.
CC TRANSFAC; T00754; -.

DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 3.
DR ProDom; PD000003; Znf_C2H2; 2.
DR SMART; SM00355; Znf_C2H2; 3.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 3.
KW Transcription regulation; Activator; Zinc-finger; Metal-binding;
KW DNA-binding; Nuclear protein; Repeat; Glycoprotein.
FT ZN FING 629 653 C2H2-TYPE 1.
FT ZN FING 659 683 C2H2-TYPE 2.
FT ZN FING 689 711 C2H2-TYPE 3.
SQ SEQUENCE 788 AA; 81015 MW; AA2B0CAB81AAB80C CRC64;

Alignment Scores:
Pred. No.: 4.6 Length: 788
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.82% Indels: 0
DB: 1 Gaps: 0

US-09-270-437D-6 (1-3412) x SP1_RAT (1-788)

QY 893 ATCTCTCGCTAGTTGGTCTCATCT 867
Db 308 IIESSerSAlaSerLeuValSerSer 316
|||||

RESULT 6

KROS_HUMAN
ID KROS_HUMAN STANDARD; PRT; 2347 AA.
AC P08922; Q15368;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Proto-oncogene tyrosine-protein kinase ROS precursor (BC 2.7.1.112)
(c-ros-1).
GN ROS1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90280463; PubMed=2352949;
RA Birchmeier C., O'Neill K., Riggs M., Wigler M.;
RT "Characterization of ROS1 cDNA from a human glioblastoma cell line.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:4799-4803(1990).
RN [2]
RP SEQUENCE OF 1790-2259 FROM N.A.
RX MEDLINE=87064611; PubMed=3023956;
RA Matsushime H., Wang L.-H., Shibuya M.;
RT "Human c-ros-1 gene homologous to the v-ros sequence of UR2 sarcoma
virus encodes for a transmembrane receptorlike molecule.";
RL Mol. Cell. Biol. 6:3000-3004(1986).
RN [3]
RP SEQUENCE OF 1854-2245 FROM N.A.
RX MEDLINE=87064625; PubMed=3785223;
RA Birchmeier C., Birnbaum D., Waitches G., Fasano O., Wigler M.;
RT "Characterization of an activated human ros gene.";
RL Mol. Cell. Biol. 6:3109-3116(1986).
CC -!- FUNCTION: THIS IS A PROBABLY A CELL GROWTH OR DIFFERENTIATION
CC FACTOR RECEPTOR WITH A TYROSINE-PROTEIN KINASE ACTIVITY.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: Belongs to the Tyr family of protein kinases. Insulin
CC receptor subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

KW Cell division; Transmembrane; Inner membrane; Coiled coil;
KW Complete proteome.
FT DOMAIN 1 3 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 4 21 POTENTIAL.
FT FT 22 PERIPLASMIC (POTENTIAL).
FT DOMAIN 27 COILED COIL (POTENTIAL).
SQ SEQUENCE 111 AA; 12658 MW; E4A926F7359C8C9C CRC64;

Alignment Scores:
Pred. No.: 56.9 Length: 111
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.73% Indels: 0
DB: 1 Gaps: 0

US-09-270-437D-6 (1-3412) x FTSB_RALSO (1-111)
QY 475 CTAATTTTCCTTCTCTCTGTTG 452
Db 6 LeupheLeuLeuLeuLeuLeu 13

RESULT 8
N551_SOYN
ID N551_SOYN STANDARD; PRT; 137 AA.
AC Q05544;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Early nodulin 55-1 precursor (N-55-1) (Fragment).
GN ENOD55-1.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Williams; TISSUE=Root;
RA MEDLINE=94003074; PubMed=8400132;
RX de Blank C., Mylona P., Katinakis P.C., Bisseling T., Franssen H.;
RT "Characterization of the soybean early nodulin cDNA clone GMEOD55."; Plant Mol. Biol. 22:1167-1171(1993).
RL -1- SUBCELLULAR LOCATION: Peribacteroid membrane (Potential).
CC -1- DEVELOPMENTAL STAGE: Expressed at early stages of nodule development. Maximal expression is seen in nodules from 14-day-old plants after which levels decrease.
CC -1- INDUCTION: During nodulation in legume roots after Rhizobium infection, and after release of bacteria from the infection thread.
CC -1- SIMILIARITY: Contains 1 plastocyanin-like domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL; X69156; CAA48908.1; -.
PIR; S37353; S37353.
DR InterPro; IPR008972; Cupredoxin.
DR Pfam; IPR003245; pcyvanin like.
DR ProDom; PD003122; Pcyvanin_like; 1.
KW Nodulation; Glycoprotein; Nitrogen fixation; Signal; Membrane.
FT NON TER 1 1
FT SIGNAL <1 ? POTENTIAL.
FT CHAIN ? 137 EARLY NODULIN 55-1.
FT DOMAIN ? ? PLASTOCYANIN-LIKE.
FT CARBOHYD 13 13 N-LINKED (GLCNAC. ..) (POTENTIAL).
FT CARBOHYD 51 51 N-LINKED (GLCNAC. ..) (POTENTIAL).
Alignment Scores:
Pred. No.: 55.4 Length: 150
Score: 8.00 Matches: 8

FT CARBOHYD 68 N-LINKED (GLCNAC. ..) (POTENTIAL).
SQ SEQUENCE 137 AA; 14857 MW; 988DE6F89A2E0BE0 CRC64;

Alignment Scores:
Pred. No.: 55.9 Length: 137
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.71% Indels: 0
DB: 1 Gaps: 0

US-09-270-437D-6 (1-3412) x N551_SOYN (1-137)
QY 2727 CCTCCCTCTCTCTCTCATCTCCA 2750
Db 78 ProSerProSerSerProSerPro 85

RESULT 9
SHUB_PIG
ID SHUB_PIG STANDARD; PRT; 150 AA.
AC P79339;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 5-hydroxytryptamine 1B receptor (5-HT-1B) (Serotonin receptor) (5-HT1B) (Fragment).
GN HTR1B.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain cortex;
RA Wurch T., Lestienne F., Colpaert F.C., Pauwels P.J.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
CC -1- FUNCTION: This is one of the several different receptors for 5-hydroxytryptamine (serotonin), a biogenic hormone that functions as a neurotransmitter, a hormone, and a mitogen. The activity of this receptor is mediated by G proteins that inhibit adenylate cyclase activity (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILIARITY: Belongs to family 1 of G-protein coupled receptors. Strongest to the other 5HT-1 subtype receptors.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL; Y11867; CAA72615.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCRRHODPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR PROSITE; PS02462; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Multigene family.
FT NON TER 1 1
FT DOMAIN <1 83 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 84 107 6 (POTENTIAL).
FT DOMAIN 108 116 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 117 141 7 (POTENTIAL).
FT DOMAIN 142 >150 CYTOPLASMIC (POTENTIAL).
FT NON TER 150 150
SQ SEQUENCE 150 AA; 16941 MW; 084EDF34A349555A CRC64;

Alignment Scores:
Pred. No.: 55.4 Length: 150
Score: 8.00 Matches: 8

Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 0.71%
DB: 1
Indels: 0
Gaps: 0
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0

US-09-270-437D-6 (1-3412) x 5H1B_PIG (1-150)

QY 2547 TCTGGATCACCTGTGTATGTCAAC 2570
|||||

DB 49 SerGlySerProValTyrValAsn 56

RESULT 10

ID 5H1B CANFA STANDARD; PRT; 161 AA.
AC P79250;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE 5-hydroxytryptamine 1B receptor (5-HT-1B) (Serotonin receptor) (5-

HT1B) (5-HT1D subtype beta) (Fragment).

GN HT1B.

OS Canis familiaris (Dog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

OX NCBI_TaxID=9615;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Beagle, and Alsatian; TISSUE=Artery;

RX MEDLINE=96296365; PubMed=8763409;

RA Sgard F., Faure C., Graham D.;

RT "Evidence for 5-HT1D beta but not 5-HT1D alpha receptor subtype

expression in canine large coronary arteries and saphenous vein.;"

RL Cardiovasc. Res. 31:793-799(1996)

CC -!- FUNCTION: This is one of the several different receptors for 5-

hydroxytryptamine (serotonin), a biogenic hormone that functions

as a neurotransmitter, a hormone, and a mitogen. The activity of

this receptor is mediated by G proteins that inhibit adenylyate

cyclase activity.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein.

CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.

CC Strongest to the other 5HT-1 subtype receptors.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; S82461; AAB37488.2; -.

DR InterPro; IPR000276; GPCR_Rhodpsn.

DR Pfam; PF00001; 7tm.1; 1.

DR PRINTS; PR00237; GPCRHHODPSN.

DR PROSITE; PS00237; G-PROTEIN RECP F1.1; PARTIAL.

DR PROSITE; PS0262; G-PROTEIN RECP F1.2; 1.

KW G-protein coupled receptor; Transmembrane; Glycoprotein;

KW Multigene family; Lipoprotein; Palmitate.

FT NON_TER 1

FT DOMAIN <1 13

FT TRANSMEM 14 37

FT DOMAIN 123 123

FT TRANSMEM 124 147

FT DOMAIN 148 156

FT TRANSMEM 157 >161

FT NON_TER 161

SQ SEQUENCE 161 AA; 17978 MW; 7BA978FA08A025B1 CRC64;

Alignment Scores:

Pred. No.:

Score: 55.1

Matches: 161

Conservative: 8

Mismatch: 0

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 0.71%
DB: 1
Indels: 0
Gaps: 0

US-09-270-437D-6 (1-3412) x 5H1B_CANFA (1-161)

QY 2547 TCTGGATCACCTGTGTATGTCAAC 2570
|||||

DB 89 SerGlySerProValTyrValAsn 96

RESULT 11

YI75_STAAM

ID YI75_STAAM STANDARD; PRT; 171 AA.

AC Q53719;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Hypothetical protein SAV1875/SA1692/MW1815 (ORF1).

GN SAV1875 OR SA1692 OR MW1815.

OS Staphylococcus aureus (strain Mu50 / ATCC 700699),

OS Staphylococcus aureus (strain N315),

OS Staphylococcus aureus (strain MW2), and

OS Staphylococcus aureus.

OC Bacteria; Firmicutes; Bacillales; Staphylococcus.

OX NCBI_TaxID=158878, 158879, 196620, 1280;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MU50 / ATCC 700699, and N315;

RX MEDLINE=21311952; PubMed=11418146;

RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,

RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,

RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,

RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,

RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,

RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,

RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;

RT "Whole genome sequencing of methicillin-resistant Staphylococcus

aureus.;"

RL Lancet 357:1225-1240(2001).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=MW2;

RX MEDLINE=22040717; PubMed=12044378;

RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,

RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,

RA Yanamoto K., Hiramatsu K.;

RT "Genome and virulence determinants of high virulence community-

acquired MRSA.;"

RL Lancet 359:1819-1827(2002).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=NCTC 8325;

RX MEDLINE=94131293; PubMed=7905453;

RA Borchardt S.A., Babwah A.V., Jayaswal R.K.;

RT "Sequence analysis of the region downstream from a peptidoglycan

hydrolase-encoding gene from Staphylococcus aureus NCTC8325.;"

RL Gene 137:253-258(1993).

CC -!- SIMILARITY: BELONGS TO THE TH1J / PPPI FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; AP003363; BAB58037.1; -.

DR EMBL; AP003363; BAB42961.1; -.

DR EMBL; AP004828; BAB95680.1; -.

DR EMBL; L19300; AAA18514.1; -.

DR PIR; B89975; B89975.

DR MEROPS; C56.UPW; -.

DR InterPro; IPR006286; Peptidase_C56.

DR InterPro: IPR002818; ThiJ/PfpI.
DR Pfam: PF01965; DJ-1_PfpI; 1.
DR TIGRams: TIGR01382; PfpI; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 171 AA; 18632 MW; A571A0890FDD719 CRC64;

Alignment Scores: 54.8 Length: 171
Pred. No.: 8.00 Matches: 8
Score: 100.00% Conservatives: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 0.71% Gaps: 0
DB: 1

US-09-270-437D-6 (1-3412) x Y175_STAAM (1-171)

QY 2527 GGAAGGACATCAGCGAGTCTG 2550

Db 119 GlyArgThrLeuThrAlaValLeu 126

RESULT 12

ID RS7_NEUCR STANDARD; PRT; 202 AA.
AC O43105;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 40S ribosomal protein S7.
GN RPS-7 OR CRPS-7.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]

SEQUENCE FROM N.A.
RC STRAIN=74-OR23-IVA;
RX MEDLINE=97174111; PubMed=9021131;
RA Vierula P.J.;
RT Cloning and characterization of a Neurospora crassa ribosomal
RT protein gene, crps-7.
RL Curr. Genet. 31:139-143(1997).
CC -1- SIMILARITY: Belongs to the S7E family of ribosomal proteins.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL: U73847; AAB94301.1; -;
DR PIR; T46586; T46586.
DR InterPro: IPR000554; Ribosomal_S7E.
DR Pfam: PF01251; Ribosomal_S7e; 1.
DR ProDom: PD006276; Ribosomal_S7E; 1.
DR PROSITE: PS00948; RIBOSOMAL_S7E; 1.
KW Ribosomal protein.
SQ SEQUENCE 202 AA; 22819 MW; 94DB7756FBE37798 CRC64;

Alignment Scores: 54 Length: 202
Pred. No.: 8.00 Matches: 8
Score: 100.00% Conservatives: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 0.71% Gaps: 0
DB: 1

US-09-270-437D-6 (1-3412) x RS7_NEUCR (1-202)

QY 1817 AGCAGGAGGAGCAATACCCCTC 1840

Db 111 SerSerArgSerArgAsnThrLeu 118

RESULT 13

YSX2 CAEEL STANDARD; PRT; 208 AA.
ID YSX2 CAEEL
AC Q10021; Q95ZN0; Q95ZN1; Q95ZN2;
DT 01-OCT-1996 (Rel. 34, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical RNA-binding protein T28D9.2 in chromosome II.
GN T28D9.2.

OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Fulton L.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [2]

REVISIONS, AND ALTERNATIVE SPLICING.

RA Waterston R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

CC -1- ALTERNATIVE PRODUCTS:

CC Event-Alternative splicing; Named isoforms=4;
CC Comment=Experimental confirmation may be lacking for some
CC isoforms;

CC Name=a;
CC IsoId=Q10021-1; Sequence=Displayed;

CC Name=b;
CC IsoId=Q10021-2; Sequence=VSP_005908, VSP_005910;

CC Name=c;
CC IsoId=Q10021-3; Sequence=VSP_005906, VSP_005907;

CC Name=d;
CC IsoId=Q10021-4; Sequence=VSP_005909;
CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL: U28738; AAG68314.2; -;
DR EMBL; U28738; AAG68399.2; -;
DR EMBL; U28738; AAG68400.1; -;
DR EMBL; U28738; AAK72066.2; -;
DR WormPep; T28D9.2a; CE29826.
DR WormPep; T28D9.2b; CE29827.
DR WormPep; T28D9.2c; CE28094.
DR WormPep; T28D9.2d; CE29828.
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF00076; rrm; 1.
DR SMART; SM00360; RRM; 1.
DR PROSITE; PS0102; RRM; 1.
DR PROSITE; PS00030; RRM_RNP_1; FALSE NEG.

KW Hypothetical protein; RNA-binding; Alternative splicing.

FT DOMAIN 2 74 RNA-BINDING (RRM).
FT DOMAIN 84 205 ASG/SER-RICH.
FT VARSPPLIC 1 128 MERLYLGKTPYNGHERDVERFLKGYKINNISKMYGFAFVD

FT FEDSRDAEDACHLDGKTWEGSMRLVEMARGKPGNDRH
FT GSRSPRRSRSPRRSRTPRRSRSRDRSRSRSRSS
FT RGRSP -> MLKTHAMIKWEKPKWALCALSWKLAENHAA
FT MTVMDRDHDHDPALLDVVHLHVEDVPALVIANVLADPVP
FT EAPDPDHQFKVVAZANPVRARKETLWNVQYIHFSP
FT SPHSYOKPLFRUKE (in isoform c).
FT /FTID=VSP_005906.
FT Missing (in isoform c).
FT /FTID=VSP_005907.
FT LKREASRSRSLPAKDRSRT -> FHSSMRNQVHLQAHIAM
FT AVT (in isoform b).

FT VARSPPLIC 129 147

FT VARSPPLIC 147 166

```
FT FT VARSPLIC 148 208 /FTid=VSP_005908.
FT KREASRSRSPAPKDRSGSPKNGGDRKRSVSRGRSH
FT SRDSNRSVSPSPGSPKSD -> QLRANISVSPSPDSD
FT SIKCKNYLIGSVKRRDVIHHCQKLEAFEDVLPORTAEIE
FT SAA (in isoform d).
FT /FTid=VSP_005909.
FT Missing (in isoform b).
FT /FTid=VSP_005910.
SQ SEQUENCE 208 AA; 23946 MW; A33CAF78D74A6F8 CRC64;

Alignment Scores:
Pred. No.: 53.9 Length: 208
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.73% Indels: 0
DB: 1 Gaps: 0

US-09-270-437d-6 (1-3412) x YSX2_CABEL (1-208)
QY 202 AGTCACGAGGCGTAGCGGACT 179
Db 93 SerProArgArgSerArgThr 100

RESULT 14
RS2_METMA
ID RS2_METMA STANDARD; PRT; 224 AA.
AC Q8PW41.
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S2P.
GN RPS2P OR MW1760.
OS Methanosarcina mazei (Methanosarcina frisia).
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2209;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Goel / Goli / ATCC BAA-199 / DSM 3647 / OCM 88;
RX MEDLINE=22120827; PubMed=12125824;
RA Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,
RA Martinez-Arias R., Henne A., Wierzer A., Baeume S., Jacobi C.,
RA Brueggemann H., Lienard T., Christmann A., Boencke M., Steckel S.,
RA Bhattacharya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
RA Fritz H.-J., Gottschalk G.
RT "The genome of Methanosarcina mazei: evidence for lateral gene
transfer between Bacteria and Archaea."
RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
CC -!- SIMILARITY: Belongs to the S2P family of ribosomal proteins.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB013412; AAM31456.1; -.
CC HAWAP; MF_00291; -.
CC InterPro; IPR001865; Ribosomal S2.
CC InterPro; IPR005707; Ribosomal S2_e/a.
CC Pfam; PF00318; Ribosomal S2_1.
CC PRINTS; PR00395; RIBOSOMALS2.
CC TIGRFAMs; TIGR01012; Sa_S2_EA; 1.
CC PROSITE; PS00962; RIBOSOMAL_S2_1; 1.
CC PROSITE; PS00963; RIBOSOMAL_S2_2; 1.
CC KX Ribosomal protein; Complete proteome.
CC SQ SEQUENCE 224 AA; 24625 MW; 144D106450EA125E CRC64;

Alignment Scores:
Pred. No.: 53.6 Length: 224
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.73% Indels: 0
DB: 1 Gaps: 0
```

```
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.73% Indels: 0
DB: 1 Gaps: 0

US-09-270-437d-6 (1-3412) x RS2_METMA (1-224)
QY 2496 CGGTGGCCCTCAAGCTCCTCTCT 2473
Db 83 ArgValAlaSerLysLeuLeuSer 90

RESULT 15
RS2_METAC
ID RS2_METAC STANDARD; PRT; 225 AA.
AC Q8TT39.
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S2P.
GN RPS2P OR MA0600.
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA FitzHugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
RA Allen N., Naylor J., Stange-Thomann N., Dearellano K., Johnson R.,
RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Guss A.M.,
RA Hedderich R., Ingram-Smith C., Kuettner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
and physiological diversity."
RL Genome Res. 12:532-542(2002).
CC -!- SIMILARITY: Belongs to the S2P family of ribosomal proteins.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB010721; AAM04044.1; ALT_INIT.
CC HAWAP; MF_00291; -.
CC InterPro; IPR001865; Ribosomal S2.
CC InterPro; IPR005707; Ribosomal S2_e/a.
CC Pfam; PF00318; Ribosomal S2_1.
CC PRINTS; PR00395; RIBOSOMALS2.
CC TIGRFAMs; TIGR01012; Sa_S2_EA; 1.
CC PROSITE; PS00962; RIBOSOMAL_S2_1; 1.
CC PROSITE; PS00963; RIBOSOMAL_S2_2; 1.
CC KX Ribosomal protein; Complete proteome.
CC SQ SEQUENCE 225 AA; 24779 MW; 02C722D26A5F411F CRC64;

Alignment Scores:
Pred. No.: 53.5 Length: 225
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.73% Indels: 0
DB: 1 Gaps: 0
```

US-09-270-437D-6 (1-3412) x RS2_METAC (1-225)

QY 2496 CCGGTGGCCTCAAAGCTCCTCTCT 2473

Db 84 ArgValAlaSerLysLeuLeuSer 91

Search completed: July 16, 2004, 11:18:44
Job time : 50 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: July 16, 2004, 10:57:48 ; Search time 162.5 Seconds

(without alignments)
13249.823 Million cell updates/sec

Title: US-09-270-437D-6

Perfect score: 1120

Sequence: 1 ggcagcggaggagcagga.....aaccttgaaatgtttattt 3412

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1017041 seqs, 315518202 residues

Word size: 1

Total number of hits satisfying chosen parameters: 2033934

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2 1/USPTO.spool_p/US09270437/runat 16072004 113127 13955/app query.fasta 1.3591
-DB=SPTREMBL 25 -QFMT=fastan -SUFFIX=Oligo.rspt -MNMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09270437 @CNG 1.1 283 @runat 16072004 113127 13955 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOF=6
-FGAPEXT=7 -YGAPOF=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

SPTREMBL 25: *
1: sp_archaea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_thc: *
8: sp_organelle: *
9: sp_phage: *
10: sp_plant: *
11: sp_rodent: *
12: sp_virus: *
13: sp_vertebrate: *
14: sp_unclassified: *
15: sp_rvirus: *
16: sp_bacteriap: *
17: sp_archaeap: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	357	31.9	556	4 Q9Y6M1	Q9y6m1 homo sapien

2	88	7.9	545	11	Q7TQF9	Q7tqf9 mus musculu
3	23	2.1	582	13	Q9PW80	Q9pw80 brachydanio
4	23	2.1	593	13	O57526	O57526 xenopus lae
5	23	2.1	594	13	O73932	O73932 xenopus lae
6	19	1.7	576	13	O42254	O42254 gallus gall
7	19	1.7	577	4	Q9NZI8	Q9nzi8 homo sapien
8	19	1.7	577	11	Q8CGX0	Q8cgx0 rattus norv
9	19	1.7	577	11	Q8BRH1	Q8brh1 mus musculu
10	19	1.7	577	11	Q80US9	Q80us9 mus musculu
11	19	1.7	577	11	O88477	O88477 mus musculu
12	19	1.7	579	4	O00425	O00425 homo sapien
13	19	1.7	579	11	Q8C2J9	Q8c2j9 mus musculu
14	19	1.7	579	11	Q9CPN8	Q9cpn8 mus musculu
15	14	1.2	169	11	Q7TP50	Q7tp50 rattus norv
16	12	1.1	100	4	Q86VB1	Q86vb1 homo sapien
17	10	0.9	215	16	Q824H3	Q824h3 chlamydomophi
18	10	0.9	1328	4	Q9UIW1	Q9uiw1 homo sapien
19	9	0.8	319	16	Q9RV01	Q9rv01 deinococcus
20	9	0.8	329	3	Q8NIY1	Q8niy1 neutrospora
21	9	0.8	366	16	Q926X1	Q926x1 listeria mo
22	9	0.8	385	11	Q9DAE2	Q9dae2 mus musculu
23	9	0.8	393	16	Q82SE0	Q82se0 nitrosomona
24	9	0.8	399	13	Q8UGN3	Q8jgn3 monoteru
25	9	0.8	422	16	Q8FLQ8	Q8flq8 corynebacte
26	9	0.8	451	5	Q18705	Q18705 caenorhabdi
27	9	0.8	512	16	Q8P646	Q8p646 xanthomonas
28	9	0.8	549	10	Q94D52	Q94d52 cryza sativ
29	9	0.8	559	11	Q8K4R0	Q8k4r0 rattus norv
30	9	0.8	623	10	Q9CAY1	Q9cay1 arabidopsis
31	9	0.8	704	5	Q9VII3	Q9vii3 drosophila
32	9	0.8	711	17	O8TVS7	O8tvs7 methanopyru
33	9	0.8	784	11	O89087	O89087 mus musculu
34	9	0.8	798	4	Q86IN8	Q86in8 homo sapien
35	9	0.8	807	10	Q9U672	Q9u672 crypthecodi
36	9	0.8	865	17	Q8Q072	Q8q072 methanosarc
37	9	0.8	893	5	Q8I429	Q8i429 plasmodium
38	9	0.8	1024	10	O9XGS7	O9xgs7 chlamydomon
39	9	0.8	1086	16	Q8EFB9	Q8efb9 shewanella
40	9	0.8	1135	10	Q7XS11	Q7xsl1 cryza sativ
41	9	0.8	1226	10	Q9LXT8	Q9lxt8 arabidopsis
42	9	0.8	1854	11	Q63131	Q63131 rattus norv
43	9	0.8	2317	11	Q63130	Q63130 rattus norv
44	9	0.8	2338	11	Q63132	Q63132 rattus norv
45	9	0.8	8625	5	Q86GD6	Q86gd6 procamburus

ALIGNMENTS

RESULT 1

Q9Y6M1 ID Q9Y6M1 PRELIMINARY; PRT; 556 AA.

AC Q9Y6M1;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Hepatocellular carcinoma autoantigen.
GN P62.

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

RX MEDLINE=99207072; PubMed=10190901;
RA Zhang J.Y., Chan E.K., Peng X.X., Tan E.M.;
RT "A novel cytoplasmic protein with RNA-binding motifs is an autoantigen
in human hepatocellular carcinoma.";
RL J. Exp. Med. 189:1101-1110(1999).
CC -!- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR EMBL; AF057352; AAD31596.1; -;
DR GO; GO:0005737; C:cytoplasm; TAS.
DR GO; GO:0003723; F:RNA binding; TAS.
DR GO; GO:0007345; P:embryogenesis and morphogenesis; TAS.

DR GO; GO:0009386; P:translational attenuation; TAS.

DR InterPro; IPR004087; KH dom.

DR InterPro; IPR004088; KH_type_1.

DR InterPro; IPR00504; RNA_rec_mot.

DR Pfam; PF00013; KH; 4.

DR Pfam; PF00076; xtm; 2.

DR SMART; SM00322; KH; 4.

DR SMART; SM00360; RRM; 2.

DR PROSITE; PS50084; KH TYPE 1; 4.

DR PROSITE; PS50102; RRM; 2.

SQ SEQUENCE 556 AA; 61842 MW; 1EDEFB100443DDC4 CRC64;

Alignment Scores:

Pred. No.: 0 Length: 556
Score: 357.00 Matches: 357
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 31.87% Indels: 0
DB: 4 Gaps: 0

US-09-270-437D-6 (1-3412) x Q9Y6M1 (1-556)

QY 70 ATGATGACACAGCTTATACATCGGGAACCTGAGACCCCGCGGTCCACGCGACGACCTCCGG 129
Db 1 MetMetAsnLysLeuTyrIleGlyAsnLeuSerProAlaValThrAlaAspAspLeuArg 20
QY 130 CAGCTCTTTGGGGACAGGAAGTGCCTCGGGGACAGAGGCTCTGAGTCCGGCTAC 189
Db 21 GlnLeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLeuLysSerGlyTyr 40
QY 190 GCCTTCGGTGACTACCCCGACAGAACCTGGGCCATCCCGGCATCGAGACCCCTCTCGGGT 249
Db 41 AlaPheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGluThrLeuSerGly 60
QY 250 AAAGTGGATTGCATGGGAAATCATGGAAGTTGATTACTCAGTCTCTAAAAGCTAAGG 309
Db 61 LysValGluLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLysLeuArg 80
QY 310 AGCAGGAAATTCAGATTTCGAAACATCCTCTCACCTGAGTGGAGGTGTGATGGA 369
Db 81 SerArgLysIleGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspGly 100
QY 370 CTTTGGTCAATATGGGACAGTGGAGATCTGGAACTGGAACAGTCAACACAGACACAAACC 429
Db 101 LeuLeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThrAspThrGluThr 120
QY 430 GCCGTGTCACAGTCACATATGCACACAGACAGAGCAAAATAGCCATCGAGAGCTA 489
Db 121 AlaValValAsnValThrTyrAlaThrArgGluGluAlaLysIleAlaMetGluLysLeu 140
QY 490 AGCGGGCATCAGTTTGAGAACTACTCCTTCAAGATTTCTACATCCCGGATGAAGAGTG 549
Db 141 SerGlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIleProAspGluVal 160
QY 550 AGTCCCGCTGGCCCTCAGCAGCCAGCGTGGGACCACTCTTCCCGGAGCAAGGC 609
Db 161 SerSerProSerProGlnArgAlaGlnArgGlyAspHisSerSerArgGluGlnGly 180
QY 610 CACGCCCTCGGGGCACTCTCAGCCAGACAGATGATTTCCCGCTCGGATCTGCTGCT 669
Db 181 HisAlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeuVal 200
QY 670 CCCACCCAGTTTGTGGTGCCATCATCGAAAGGAGGCTTGACCAATAAGAAACATCACT 729
Db 201 ProThrGlnPheValGlyAlaIleIleGlyLysGluGlyLeuThrIleLysAsnIleThr 220
QY 730 AAGCAGACCCAGTCCCGGTAGATATCCATGAAAGAGAACTCTGAGCTGCGAGAAG 789
Db 221 LysGlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGlyAlaAlaGluLys 240
QY 790 CCTGTCACCATCCATGCCACCCAGAGGGGACTTCTGAAGCATGCCGCATGATCTTGA 849
Db 241 ProValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeuGlu 260

QY 850 ATCATGCAGAAAGAGCGAGATGAGACCAAACTAGCCGAGAGATTCCTCTGAAAAATCTTG 909
Db 261 IleMetGlnLysGluAlaAspGluThrLysLeuAlaGluGluIleProLeuLysIleLeu 280
QY 910 GCACACAATGGCTTGGTGGAAAGACTGATTCGAAAGAGGACAGAAATTTGAAGAAAT 969
Db 281 AlaHisAsnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLysIle 300
QY 970 GAACATGAACAGGACCAAGATACAAATCTCATCTTTTCAGAGGATTTGAGCATATACAC 1029
Db 301 GluHisGluThrGlyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsn 320
QY 1030 CCGGAAGAACCATCCTGTCGAGGGCACAGTTGAGGCTGTGCCAGTCTGAGATAGAG 1089
Db 321 ProGluArgThrIleThrValLysGlyThrValGluAlaCysAlaSerAlaGluIleGlu 340
QY 1090 ATTATGAAGAAAGCTCGCTGAGGCCCTTTGAAAATGATATGCTGGCTGTAAAC 1140
Db 341 IleMetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaValAsn 357

RESULT 2

Q7TQF9

AC Q7TQF9 PRELIMINARY; PRT; 545 AA.

DT 01-OCT-2003 (TrEMBLrel. 25, Created)

DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Hypothetical protein.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6; TISSUE=Brain;

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.W., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs S.W.,

RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,

RA Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

and mouse cDNA sequences."

RL proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6; TISSUE=Brain;

RA Strausberg R.;

RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.

RL EMBL; BC054552; AAH54552.1; -.

KW Hypothetical protein.

SQ SEQUENCE 545 AA; 59679 MW; 88DD35FF30DF3091 CRC64;

Alignment Scores:

Pred. No.: 2,13e-82 Length: 545
Score: 88.00 Matches: 88
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 7.86% Indels: 0
DB: 11 Gaps: 0


```

DR SMART; SM00322; KH; 4.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS0084; KH TYPE 1; 4.
DR PROSITE; PS0102; RRM; 2.
DR PROSITE; PS0030; RRM_RNP_1; FALSE NEG.
SQ SEQUENCE 593 AA; 65385 MW; 5A5AB4B4A1D55DF7 CRC64;

Alignment Scores:
Pred. No.: 3.75e-14 Length: 593
Score: 23.00 Matches: 23
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.05% Indels: 0
DB: 13 Gaps: 0

US-09-270-437D-6 (1-3412) x O57526 (1-593)

QY 1630 GCTGCGCGGTGATTGGCAAGAGTGGCAAGACCGTGAACGAACTCGAAGCTTAACCAAGT 1689
|||||
Db 511 AlaGlyArgValIleGlyLysThrValAsnGluLeuGlnAsnLeuThrSer 530

QY 1690 GCAGAAAGTC 1698
|||||
Db 531 AlaGluVal 533

RESULT 5
O73932 PRELIMINARY; PRT; 594 AA.
AC O73932
DT 01-AUG-1998 (TRENBLrel. 07, Created)
DT 01-AUG-1998 (TRENBLrel. 07, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Vg1 RNA binding protein variant D.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=83355;
RN [1]
RP SEQUENCE FROM N.A.
RA Havin L., Git A., Elisha Z., Oberman F., Yaniv K.,
RA Pressman Schwartz S., Standart N.M., Yisraeli J.K.;
RL Genes Dev. 0:0-0(1998).
RW [2]
SQ SEQUENCE FROM N.A.
RX MEDLINE=98228351; PubMed=9560341;
RA Deshler J.O., Highett M.I., Abramson T., Schnapp B.J.;
RT "A highly conserved RNA-binding protein for cytoplasmic mRNA
RT localization in vertebrates.";
RL Curr. Biol. 8:489-496(1998).
CC -|- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR EMBL; AF064634; AAC18598.1; -
DR EMBL; AF055923; AAC41285.1; -
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR004087; KH dom.
DR InterPro; IPR004088; KH type 1.
DR Pfam; PF00013; KH; 4.
DR SMART; SM00322; KH; 4.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS0084; KH TYPE 1; 4.
DR PROSITE; PS0102; RRM; 2.
DR PROSITE; PS0030; RRM_RNP_1; FALSE NEG.
SQ SEQUENCE 594 AA; 65643 MW; 54CEA7BFF0856DD6 CRC64;

Alignment Scores:
Pred. No.: 3.75e-14 Length: 594
Score: 23.00 Matches: 23
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.05% Indels: 0
DB: 13 Gaps: 0

```

```

US-09-270-437D-6 (1-3412) x O73932 (1-594)

QY 1630 GCTGCGCGGTGATTGGCAAGAGTGGCAAGACCGTGAACGAACTCGAAGCTTAACCAAGT 1689
|||||
Db 512 AlaGlyArgValIleGlyLysThrValAsnGluLeuGlnAsnLeuThrSer 531

QY 1690 GCAGAAAGTC 1698
|||||
Db 532 AlaGluVal 534

RESULT 6
O42254 PRELIMINARY; PRT; 576 AA.
AC O42254
DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Zipcode-binding protein.
OS ZBPI.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97220007; PubMed=9121465;
RA Ross A.F., Oleynikov Y.S., Kisilevsky E.H., Taneja K.L., Singer R.H.;
RT "Characterization of a beta-actin mRNA zipcode-binding protein.";
RL Mol. Cell. Biol. 17:2158-2165(1997).
CC -|- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR EMBL; AF026527; AAB82295.1; -
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR004087; KH dom.
DR InterPro; IPR004088; KH type 1.
DR InterPro; IPR00504; RNA_rec_mot.
DR Pfam; PF00013; KH; 4.
DR Pfam; PF00076; rrm; 2.
DR SMART; SM00322; KH; 4.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS0084; KH TYPE 1; 4.
DR PROSITE; PS0102; RRM; 2.
DR PROSITE; PS0030; RRM_RNP_1; FALSE NEG.
SQ SEQUENCE 576 AA; 63271 MW; 01AFA2D1D81C8811 CRC64;

Alignment Scores:
Pred. No.: 5.97e-10 Length: 576
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.70% Indels: 0
DB: 13 Gaps: 0

US-09-270-437D-6 (1-3412) x O42254 (1-576)

QY 1630 GCTGCGCGGTGATTGGCAAGAGTGGCAAGACCGTGAACGAACTCGAAGCTTAACCAAGT 1686
|||||
Db 498 AlaGlyArgValIleGlyLysThrValAsnGluLeuGlnAsnLeuThr 516

RESULT 7
Q9NZI8 PRELIMINARY; PRT; 577 AA.
AC Q9NZI8
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE mRNA-binding protein CRDBP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RA Ioannidis P., Trangas T., Dimitriadis E., Samiotaki M.,
RA Panoutsakopoulos G., Kyriazoglou I., Voutzoulis S., Tsiapalis C.M.,
RA Kittas C., Agnantis N., Pandis N.;
RT "Ectopic expression of a KH-domain containing protein, highly
RT homologous to both human IMP-1 and mouse CRD-BP, in benign and
RT malignant mesenchymal tumors.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR EMBL; AF198254; AAF37203.1; -.
DR HSSP; P11940; 1CVJ.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR004087; KH dom.
DR InterPro; IPR004088; KH_type_1.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00013; KH; 4.
DR Pfam; PF00076; rrm; 2.
DR SMART; SM00322; KH; 4.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50084; KH_TYPE_1; 4.
DR PROSITE; PS50102; RRM; 2.
SQ SEQUENCE 577 AA; 63456 MW; 0749A060F252D81D CRC64;

Alignment Scores:
Pred. No.: 5.97e-10 Length: 577
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.70% Indels: 0
DB: 4 Gaps: 0

US-09-270-437D-6 (1-3412) x Q3NZI8 (1-577)

QY 1489 AGATGGTCATCATCCGGCCACCGAAGCCAGTTCAGCCGACGGACGGATC 1545
Db 452 ArgMetValIleIlethrGlyProProGluAlaGlnPhelysAlaGlnGlyArgIle 470

RESULT 8
Q8CGX0
ID Q8CGX0 PRELIMINARY; PRT; 577 AA.
AC Q8CGX0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE B-actin zipcode binding protein 1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA Eom T., Singer R.H., Bassell G.J.;
RT "Molecular interactions between rZBP1 and b-actin zipcode required for
RT transport of mRNA and stimulation of spine growth.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF541940; AA016210.1; -.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR004087; KH dom.
DR InterPro; IPR004088; KH_type_1.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00013; KH; 4.
DR Pfam; PF00076; rrm; 2.
DR SMART; SM00322; KH; 4.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50084; KH_TYPE_1; 4.
DR PROSITE; PS50102; RRM; 2.
SQ SEQUENCE 577 AA; 63436 MW; 0647676128FBD1EE CRC64;

Alignment Scores:
Pred. No.: 5.97e-10 Length: 577
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.70% Indels: 0
DB: 4 Gaps: 0

```

```

Query Match: 1.70% Indels: 0
DB: 11 Gaps: 0

US-09-270-437D-6 (1-3412) x Q8CGX0 (1-577)

QY 1630 GCTGGCCGGTGTGTTGGCAAGGTGCAAGCCGTCGACGAACTGCAGAACTTAACC 1686
Db 499 AlaGlyArgValIleGlyLysThrValAsnGluLeuGlnAsnLeuThr 517

RESULT 9
Q8BRH1
ID Q8BRH1 PRELIMINARY; PRT; 577 AA.
AC Q8BRH1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Insulin-like growth factor 2.
GN IGF2BP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK044850; BAC32119.1; -.
DR MGD; MGI:1890357; IGF2bp1.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR004087; KH dom.
DR InterPro; IPR004088; KH_type_1.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00013; KH; 4.
DR Pfam; PF00076; rrm; 2.
DR SMART; SM00322; KH; 4.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50084; KH_TYPE_1; 4.
DR PROSITE; PS50102; RRM; 2.
SQ SEQUENCE 577 AA; 63378 MW; D75897450841532E CRC64;

Alignment Scores:
Pred. No.: 5.97e-10 Length: 577
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.70% Indels: 0
DB: 11 Gaps: 0

US-09-270-437D-6 (1-3412) x Q8BRH1 (1-577)

QY 1630 GCTGGCCGGTGTGTTGGCAAGGTGCAAGCCGTCGACGAACTGCAGAACTTAACC 1686
Db 499 AlaGlyArgValIleGlyLysThrValAsnGluLeuGlnAsnLeuThr 517

RESULT 10
Q80US9
ID Q80US9 PRELIMINARY; PRT; 577 AA.
AC Q80US9;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Igf2bp1 protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=129/Sv x 129/Sv-CP; TISSUE=Embryonic stem cells;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.W., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Faxmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Wozney K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzyzinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv x 129/Sv-CP; TISSUE=Embryonic stem cells;
RA Strausberg R.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC051679; AAH51679.1; -
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR004087; KH dom.
DR InterPro; IPR004088; KH type_1.
DR Pfam; PF00013; KH; 4.
DR Pfam; PF00076; rrm; 2.
DR SMART; SM00322; KH; 4.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50084; KH TYPE_1; 4.
DR PROSITE; PS50102; RRM; 2.
SQ SEQUENCE 577 AA; 63378 MW; D439F544257DA3CE CRC64;

Alignment Scores:
Pred. No.: 5.97e-10 Length: 577
Percent: 19.00 Matches: 19
Score: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.70% Indels: 0
DB: 11 Gaps: 0

US-09-270-437D-6 (1-3412) x Q80US9 (1-577)
Qy 1630 GCTGCGCGGTTGGCAAGTGGCAAGCCGTGAACGAACTGCAGAACTTAACC 1686
Db 499 AlAGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsnLeuThr 517

RESULT 11
O88477
ID O88477 PRELIMINARY; PRT; 577 AA.
AC
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Coding region determinant binding protein.
GN IGF2BP1 OR CRDBP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBT_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92217743; PubMed=1559612;
RA Bernstein P.L., Herrick D.J., Prokipcak R.D., Ross J.;
RA "Control of c-myc mRNA half-life in vitro by a protein capable of
RT binding to a coding region stability determinant.";

```

```

RL Genes Dev. 6:642-654 (1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94158886; PubMed=8111742;
RA Herrick D.J., Ross J.;
RT "The half-life of c-myc mRNA in growing and serum-stimulated cells:
RT influence of the coding and 3' untranslated regions and role of
RT ribosome translocation.";
RL Mol. Cell. Biol. 14:2119-2128 (1994).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=94179348; PubMed=8132663;
RA Prokipcak R.D., Herrick D.J., Ross J.;
RA "Purification and properties of a protein that binds to the C-terminal
RT coding region of human c-myc mRNA.";
RL J. Biol. Chem. 269:9261-9269 (1994).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=9732234; PubMed=9178888;
RA Leeds P., Kren B.T., Boylan J.M., Betz N.A., Steer C.J.,
RA Gruppiso P.A., Ross J.;
RT "Developmental regulation of CRD-BP, an RNA-binding protein that
RT stabilizes c-myc mRNA in vitro.";
RL Oncogene 14:1279-1286 (1997).
RN [5]
RP SEQUENCE FROM N.A.
RX Ross J., Prokipcak R.D., Leeds P., Doyle G.A.R., Betz N.A.,
RA Fleisig A.J.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamakawa I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikolaïdo I., Pesole G., Quackenbush J.,
RA Schram L.M., Straubli P., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaurts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).
CC -!- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR EMBL; AF061569; AAC72743.1; -.
DR EMBL; AK013940; BAB29071.1; -.
DR HSP; P11940; 1CVJ.
DR MGD; MGI:1890357; Igfbp1.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR004087; KH dom.
DR InterPro; IPR004088; KH type_1.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00013; KH; 4.
DR Pfam; PF00076; rrm; 2.
DR SMART; SM00322; KH; 4.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50084; KH TYPE_1; 4.
DR PROSITE; PS50102; RRM; 2.
SQ SEQUENCE 577 AA; 63450 MW; EPBBLAF2FF9F0344 CRC64;

Alignment Scores:
Pred. No.: 5.97e-10 Length: 577
Score: 19.00 Matches: 19

```

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.70% Indels: 0
DB: 11 Gaps: 0

US-09-270-437D-6 (1-3412) x 088477 (1-577)

QY 1630 GCTGGCGGGTGAATTCGCAAGTGGCAGACCGTGAACGCTGAGAACTGAGAACTTAACC 1686

Db 499 AlaglyArgValileGlyLysGlyGlyLysThrValAsnGluLeuGlnAsnLeuThr 517

RESULT 12

ID 000425 PRELIMINARY; PRT; 579 AA.

AC 000425;

DT 01-JUL-1997 (TrEMBLrel. 04, Created)

DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Putative RNA binding protein KOC (KOC).

GN KOC.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Mueller-Pillasch F., Lacher U., Wallrapp C., Et AL.;

RL Oncogene 0:0-0(0).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Pancreas;

RA Mueller-Pillasch F., Lacher U., Wallrapp C.;

RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.

CC -!- SIMILARITY: CONTAINS 4 KH DOMAINS.

DR EMBL; U97188; AAC35208.1; --

DR EMBL; U76705; AAD09223.1; --

DR GO; GO:0005737; Cytoplasm; TAS.

DR GO; GO:0003723; rRNA binding; TAS.

DR GO; GO:0007345; P:proteinogenesis and morphogenesis; TAS.

DR GO; GO:0006412; P:protein biosynthesis; TAS.

DR GO; GO:0006396; P:RNA processing; TAS.

DR InterPro; IPR004087; KH dom.

DR InterPro; IPR004088; KH_type_1.

DR InterPro; IPR000504; RNA_rec_mot.

DR Pfam; PF00013; KH; 4.

DR Pfam; PF00076; xrm; 2.

DR SMART; SM00322; KH; 4.

DR SMART; SM00360; RRM; 2.

DR PROSITE; PS50084; KH_TYPE_1; 4.

DR PROSITE; PS50102; RRM; 2.

DR PROSITE; PS00030; RRM_RNP_1; FALSE NEG.

SQ SEQUENCE 579 AA; 63720 MW; AE53A8EE3C135C5 CRC64;

Alignment Scores:

Pred. No.: 5 97e-10 Length: 579
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.70% Indels: 0
DB: 4 Gaps: 0

US-09-270-437D-6 (1-3412) x 000425 (1-579)

QY 1489 AGGATGTCATCATCCGGCCACCGGACCGGACCGGACCGGACCGGATC 1545

Db 452 ArgMetVallelleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 470

RESULT 13

Q8C2J9

ID Q8C2J9 PRELIMINARY; PRT; 579 AA.

AC Q8C2J9;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Insulin-like growth factor 2.

GN IGF2BP3.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=NOD; TISSUE=Thymus;

RX MEDLINE=22354683; PubMed=12466851;

RA The FANTOM Consortium.

RA the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of

RL Nature 420:563-573(2002)."

DR EMBL; AK088465; BAC40370.1; --

DR MGD; MGI:1890359; Igf2bp3.

DR GO; GO:0003676; F:nucleic acid binding; IEA.

DR InterPro; IPR004087; KH_dom.

DR InterPro; IPR004088; KH_type_1.

DR InterPro; IPR000504; RNA_rec_mot.

DR Pfam; PF00013; KH; 4.

DR Pfam; PF00076; xrm; 2.

DR SMART; SM00322; KH; 4.

DR SMART; SM00360; RRM; 2.

DR PROSITE; PS50084; KH_TYPE_1; 4.

DR PROSITE; PS50102; RRM; 2.

SQ SEQUENCE 579 AA; 63551 MW; 937B601A95D06B77 CRC64;

Alignment Scores:

Pred. No.: 5 97e-10 Length: 579
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.70% Indels: 0
DB: 11 Gaps: 0

US-09-270-437D-6 (1-3412) x Q8C2J9 (1-579)

QY 1489 AGGATGTCATCATCCGGCCACCGGACCGGACCGGACCGGACCGGATC 1545

Db 452 ArgMetVallelleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIle 470

RESULT 14

Q9CFN8

ID Q9CFN8 PRELIMINARY; PRT; 579 AA.

AC Q9CFN8;

DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE 10 days embryo cDNA, RIKEN full-length enriched library,

DE clone:2610036B18, full insert sequence (Igf2 mRNA-binding protein 3)

DE (Insulin-like growth factor 2, binding protein 3).

GN IGF2BP3 OR 2610101N1RIK OR MIMP3.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Embryo;

RX MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shingagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,

RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,

RA Kato K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,

RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,

RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,

RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombearts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Mori H., Sakakibara S., Imai T., Nakamura Y., Iijima T., Suzuki A.,
RA Yuasa Y., Takeda M., Okano H.;
RT "Expression of mouse igf2 mRNA-binding protein 3 and its implications
RT for the developing central nervous system.";
RL J. Neurosci. Res. 0:0-0(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain, and Olfactory epithelium;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.O., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzly D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain, and Olfactory epithelium;
RA Strausberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 4 KH DOMAINS.
DR EMBL; AK011689; BAB27779.1; -;
DR EMBL; AB046173; BAB19755.1; -;
DR EMBL; BC045138; AAH45138.1; -;
DR EMBL; BC049082; AAH49082.1; -;
DR MGD; MGI:1890359; Igf2bp3.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR004087; KH_dom.
DR InterPro; IPR004088; KH_type_1.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00013; KH; 4.
DR Pfam; PF00076; rim; 2.
DR SMART; SM00322; KH; 4.
DR SMART; SM00360; RRM; 2.
DR PROSITE; PS50084; KH_TYPE_1; 4.
DR PROSITE; PS50102; RRM; 2.
SQ SEQUENCE 579 AA; 63574 MW; CABD9A435B392B7 CRC64;

Alignment Scores:
Pred. No.: 5, 97e-10 Length: 579
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.70% Indels: 0
DB: 11 Gaps: 0

US-09-270-437D-6 (1-3412) x Q9CPN8 (1-579)

QY 1489 AGGATGGTCATCATCACCGGCCACCGAGCCAGTTCAGGCCAGCGGATC 1545
|||||
DB 452 ArgMetValIleIleThrGlyProGluAlaGlnPhelysAlaGlnGlyArgile 470
|||||
RESULT 15
Q7TP50 PRELIMINARY; PRT; 169 AA.
ID Q7TP50
AC Q7TP50;
DT 01-OCT-2003 (Tremblrel. 25, Created)
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
DE 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Ab2-255.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Xu C.S., Li W.Q., Li Y.C., Yang K.J., Yan H.M., Chang C.F., Zhao L.F.,
RA Ma H., Wang L., Wang S.F., Han H.P., Wang G.P., Chai L.Q., Yuan J.Y.,
RA Shi J.B., Rahman S., Wang Q.N., Zhang J.B.;
RT "Liver regeneration after PH.";
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY325199; AAP92600.1; -;
SQ SEQUENCE 169 AA; 17743 MW; 19BF6295C1000CBA CRC64;

Alignment Scores:
Pred. No.: 0.000123 Length: 169
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.25% Indels: 0
DB: 11 Gaps: 0

US-09-270-437D-6 (1-3412) x Q7TP50 (1-169)

QY 1489 AGGATGGTCATCATCACCGGCCACCGAGCCAGTTCAG 1530
|||||
DB 61 ArgMetValIleIleThrGlyProGluAlaGlnPhelys 74
|||||

Search completed: July 16, 2004, 11:29:33
Job time : 183.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 21, 2004, 12:27:22 ; Search time 8233 Seconds
(without alignments)
12375.776 Million cell updates/sec

Title: US-09-270-437D-6
Perfect score: 3412
Sequence: 1 ggcacggaggaggcgagga.....aaccttgaaaatgtttattt 3412

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 0

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_man:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vri:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	834	24.4	874	13 BQ691878	BQ691878 AGENCOURT
2	815	23.9	905	12 BM453327	BM453327 AGENCOURT
3	771	22.6	832	13 BX644668	BX644668 DKFZp781G
4	747	21.9	851	13 BQ221568	BQ221568 AGENCOURT

5	734	21.5	940	14	CD243351	AGENCOURT
6	715	21.0	824	14	CF593505	AGENCOURT
7	702	20.6	1089	12	BM806132	AGENCOURT
8	698	20.5	999	12	BM561242	AGENCOURT
9	670	19.6	724	12	BM977219	UI-CF-DU1
10	642	18.8	836	13	BQ212500	AGENCOURT
11	634	18.6	647	13	BU674592	UI-CF-DU0
12	617	18.1	889	12	BG748346	AGENCOURT
13	594	17.4	781	12	BG575889	AGENCOURT
14	590	17.3	894	13	BU156245	AGENCOURT
15	585	17.1	918	13	BQ212353	AGENCOURT
16	584	17.1	584	13	BX642777	DKFZp781I
17	566	16.6	762	13	BU601634	AGENCOURT
18	535	15.4	628	12	BI862502	AGENCOURT
19	507	14.9	939	13	BQ961314	AGENCOURT
20	505	14.8	607	12	BM724394	UI-E-E01
21	489	14.3	612	13	BG679082	UI-CF-DU1
22	480	14.1	531	10	BG009869	QV1-GN032
23	477	14.0	620	10	BF914566	IL3-UT011
24	448	13.1	729	10	BE278075	AGENCOURT
25	440	12.9	918	14	CF995222	AGENCOURT
26	440	12.9	981	13	BQ887769	AGENCOURT
27	438	12.8	453	9	AL121466	DKFZp762L
28	434	12.7	594	12	BI858380	AGENCOURT
29	423	12.4	459	10	BE277048	AGENCOURT
30	416	12.2	638	12	BM843467	K-EST0121
31	409	12.0	449	13	BQ361219	PMO-OT023
32	397	11.6	485	13	BQ355956	PMO-HT119
33	397	11.6	485	13	BQ355967	PMO-HT116
34	380	11.1	759	10	BE622021	601440673
35	378	11.1	720	13	BX101497	AGENCOURT
36	376	11.0	976	10	BG025065	AGENCOURT
37	374	11.0	386	9	AV698363	AGENCOURT
38	365	10.7	845	13	BU155790	AGENCOURT
39	357	10.5	469	13	EX282561	AGENCOURT
40	346	10.1	1212	13	BU902212	AGENCOURT
41	343	10.1	344	9	AI004702	AGENCOURT
42	341	10.0	506	10	BE928277	AGENCOURT
43	326	9.6	1078	12	BM462051	AGENCOURT
44	325	9.5	526	10	BE96980	AGENCOURT
45	310	9.1	608	12	BG254615	AGENCOURT

ALIGNMENTS

RESULT 1
BQ691878
LOCUS
DEFINITION
AGENCOURT_8034698 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6207026
5', mRNA sequence.
BQ691878
BQ691878.1 GI:21817194
EST.
KEYWORDS
SOURCE
Homo sapiens (human)
Homo sapiens
Rukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 874)
NTH-MGC <http://mgc.nci.nih.gov/>
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE
Unpublished (1999)
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: L1CM2362 row: h column: 03
High quality sequence stop: 653.

```
FEATURES
source
Location/Qualifiers
1..874
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6207026"
/tissue_type="ductal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 110"
/note="Organ: pancreas; Vector: pOTB7; Site:1: XhoI;
Site:2: EcoRI; cDNA made by oligo-dr priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 24.4%; Score 834; DB 13; Length 874;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 834; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1824 GGAGCAGAAATACCTCAGGAGTGCCTCAGCGCAGCAAGTGAAGTCCACAGCA 1883
DB 1 GGAGCAGAAATACCTCAGGAGTGCCTCAGCGCAGCAAGTGAAGTCCACAGCA 60

QY 1884 CCAGCAAAACAGCATGAATAGTCTTCCACACCTGACAGATGAGACCAACGCA 1943
DB 61 CCAGCAAAACAGCATGAATAGTCTTCCACACCTGACAGATGAGACCAACGCA 120

QY 1944 GCCAGCCAGATCGGAGCAAAACAAAGACCATCTGAGGAATGAGAAGTCTCGGAGGCG 2003
DB 121 GCCAGCCAGATCGGAGCAAAACAAAGACCATCTGAGGAATGAGAAGTCTCGGAGGCG 180

QY 2004 CCAGGACCTCTGCCAGGCCCTGAGAACCCAGGGCCGAGAGGGGGGGAAGTCA 2063
DB 181 CCAGGACCTCTGCCAGGCCCTGAGAACCCAGGGCCGAGAGGGGGGGAAGTCA 240

QY 2064 CCAGGTTTCCAGAACCAACAGGCCCGCCCTCCGCCGCCCGAGGCTTCTCAGGCTCA 2123
DB 241 CCAGGTTTCCAGAACCAACAGGCCCGCCCTCCGCCGCCCGAGGCTTCTCAGGCTCA 300

QY 2124 GCCATCCACTTCAACATCCACTCGGATCTCTCTGAATCCCAAGCAGCTATCCCTTTTA 2183
DB 301 GCCATCCACTTCAACATCCACTCGGATCTCTCTGAATCCCAAGCAGCTATCCCTTTTA 360

QY 2184 GTTGAACCTAAACATAGTGAAGTGTCAAGCCAGCAAGCAAAATGCACACCTTTTCTGTG 2243
DB 361 GTTGAACCTAAACATAGTGAAGTGTCAAGCCAGCAAGCAAAATGCACACCTTTTCTGTG 420

QY 2244 GCAAAATCGTCTGTACATGTGTGTACATATTAGAAAGGGAAGATGTTAAGATATGTGGC 2303
DB 421 GCAAAATCGTCTGTACATGTGTGTACATATTAGAAAGGGAAGATGTTAAGATATGTGGC 480

QY 2304 CTGTGGGTTACACAGGTCCTGACGGTAAATATATTTTGAATAATATATCAATAA 2363
DB 481 CTGTGGGTTACACAGGTCCTGACGGTAAATATATTTTGAATAATATATCAATAA 540

QY 2364 CTCAACTAACCTCAATTTTAAATCAATTAATTTTCTTTTCTTTTAAAGAGCA 2423
DB 541 CTCAACTAACCTCAATTTTAAATCAATTAATTTTCTTTTCTTTTAAAGAGCA 600

QY 2424 GGCCTTTCTAGACTTTAAGATAAAGTCTTTTGGGAGGTCTCAGCGTGTAGAGAGGCT 2483
DB 601 GGCCTTTCTAGACTTTAAGATAAAGTCTTTTGGGAGGTCTCAGCGTGTAGAGAGGCT 660

QY 2484 TTGAGGCCACCCGACAAAATTCACCCAGAGGAAATCTCGTCCGAGGACACTCAACGGC 2543
DB 661 TTGAGGCCACCCGACAAAATTCACCCAGAGGAAATCTCGTCCGAGGACACTCAACGGC 720

QY 2544 AGTTCTGATCACTGTGTATGTCAACAGAGGATACCGTCTCTCTGAGAGGAATCTC 2603
DB 2544 AGTTCTGATCACTGTGTATGTGTCAACAGAGGATACCGTCTCTCTGAGAGGAATCTC 2603

Db 721 AGTTCTGATCACTGTGTATGTGTCAACAGAGGATACCGTCTCTCTGAGAGGAATCTC 780
QY 2604 TGTCACTCTCTATGCTCTGTAGCTCATACCCCATTTCTCTCTCTCTCTCTCTCTC 2657
DB 781 TGTCACTCTCTATGCTCTGTAGCTCATACCCCATTTCTCTCTCTCTCTCTCTCTC 834

RESULT 2
BM453327 905 bp mRNA linear EST 05-FEB-2002
LOCUS AGENCOURT_6397801 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5527679
5', mRNA sequence.
ACCESSION BM453327
VERSION BM453327.1 GI:18502367
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 905)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12203 Row: e Column: 24
High quality sequence stop: 568.

FEATURES
source
Location/Qualifiers
1..905
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5527679"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 72"
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 2 kb. Library constructed by Life
Technologies."
```

```
QY 792 TGTCACCATCCATCCACCCAGAGGGGACTTCTGAAGCATGCCGATGATCTTGAAAT 851
Db 301 TGTCACCATCCATCCACCCAGAGGGGACTTCTGAAGCATGCCGATGATCTTGAAAT 360
QY 852 CATGCAGAAAGAGGCGAGATGAGACCAAACTAGCCGAAGAGATTCCTCTGAAATCTTTGGC 911
Db 361 CATGCAGAAAGAGGCGAGATGAGACCAAACTAGCCGAAGAGATTCCTCTGAAATCTTTGGC 420
QY 912 ACACAATGCTGCTGTTGGAAGACTGATGGAAAGAGGAGGAGCAAAATTTGAAGAAATTTGA 971
Db 421 ACACAATGCTGCTGTTGGAAGACTGATGGAAAGAGGAGGAGCAAAATTTGAAGAAATTTGA 480
QY 972 ACATGAAACAGGACCAAGATACAACTCTCATCTTTTGAGGATTTGAGCATATACAAACC 1031
Db 481 ACATGAAACAGGACCAAGATACAACTCTCATCTTTTGAGGATTTGAGCATATACAAACC 540
QY 1032 GGAAGAACCATCACTGTGAAGGGCACAGTTGAGGCTGTGCCAGTCTGAGATAGAGAT 1091
Db 541 GGAAGAACCATCACTGTGAAGGGCACAGTTGAGGCTGTGCCAGTCTGAGATAGAGAT 600
QY 1092 TATGAAGAGCTCGTGAGGCTTTGAAATGATATGCTGGCTGTAAACCAAGCCAA 1151
Db 601 TATGAAGAGCTCGTGAGGCTTTGAAATGATATGCTGGCTGTAAACCAAGCCAA 660
QY 1152 TCTGATCCAGGGTTGAACCTCAGCGCACCTTGGCATCTTTTCAACAGGACTGTCCGTGCT 1211
Db 661 TCTGATCCAGGGTTGAACCTCAGCGCACCTTGGCATCTTTTCAACAGGACTGTCCGTGCT 720
QY 1212 ATCTCCACGACGAGGCCCCGGGAGCTCCCCCGCTGTCGCCCTTACCACCCCTTCACTAC 1271
Db 721 ATCTCCACGACGAGGCCCCGGGAGCTCCCCCGCTGTCGCCCTTACCACCCCTTCACTAC 780
QY 1272 CCACTCGGATCTTCTCCAGCCTGTACCCCATC 1306
Db 781 CCACTCGGATCTTCTCCAGCCTGTACCCCATC 815
```

```
RESULT 3
BX644668
LOCUS
DEFINITION
DKFZP781G0234 r1 781 (synonym: hlcc4) Homo sapiens cDNA clone
DKFZP781G0234 5', mRNA sequence.
```

```
ACCESSION
BX644668
VERSION
BX644668.1 GI:34479001
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
1 (bases 1 to 832)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Lauber,J., Bahr,A., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
Fobo,G., Han,M. and Wiemann,S.
EST (Lauber,J., Bahr,A., Mewes,H.W., Weil,B., Amid,C., et al.)
Unpublished (2003)
Contact: MIPS
MIPS
```

```
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
consortium of the German Genome Project.
No sl sequence available.
This clone (DKFZP781G0234) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1..832
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZP781G0234"
/dev_stage="adult"
```

FEATURES

source

```
ORIGIN
Query Match 22.6%; Score 771; DB 13; Length 832;
Best Local Similarity 100.0%; Pred. No. 1.9e-294; Indels 0; Gaps 0;
Matches 771; Conservative 0; Mismatches 0;
/lab_host="DH10B"
/clone_lib="781 (synonym: hlcc4)"
/note="vector: pSport1_Sfi; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"
QY 1732 GAGGAAGTATCGTCAGAAATTTATCGGGCAGCTTTCTTTGCTAGCCAGACTGCACAGCGCAAG 1791
Db 13 GAGGAAGTATCGTCAGAAATTTATCGGGCAGCTTTCTTTGCTAGCCAGACTGCACAGCGCAAG 72
QY 1792 ATCAGGGAATTTGTACAAACAGGTGAAGCAGGAGCAGAAATACCTCAGGAGTGCCTC 1851
Db 73 ATCAGGGAATTTGTACAAACAGGTGAAGCAGGAGCAGAAATACCTCAGGAGTGCCTC 132
QY 1852 TCACAGCGCAGCAAGTGAGGCTCCACAGGCAACAGCAAAACAAACGATGAATGTAGCCC 1911
Db 133 TCACAGCGCAGCAAGTGAGGCTCCACAGGCAACAGCAAAACAAACGATGAATGTAGCCC 192
QY 1912 TTCCAAACACCTGACAGAATGAGACCAACGACCCAGCAGATCGGGAGCAAAACCAAGA 1971
Db 193 TTCCAAACACCTGACAGAATGAGACCAACGACCCAGCAGATCGGGAGCAAAACCAAGA 252
QY 1972 CCATCTGAGGAATGAGAACTCTGCGAGGCGGCGAGGACTCTGCGAGGCGCTTGAGAAC 2031
Db 253 CCATCTGAGGAATGAGAACTCTGCGAGGCGGCGAGGACTCTGCGAGGCGCTTGAGAAC 312
QY 2032 CCACAGGGCGGAGGCGGCGGAGGTCAGCAGGTTTGCAGAACCAACCGAGGCCCG 2091
Db 313 CCACAGGGCGGAGGCGGCGGAGGTCAGCAGGTTTGCAGAACCAACCGAGGCCCG 372
QY 2092 CTCTCCGCGCCCGCAGGCGCTTCTGAGGCTTTCAGCCATCCACTTCCACTCGGATC 2151
Db 373 CTCTCCGCGCCCGCAGGCGCTTCTGAGGCTTTCAGCCATCCACTTCCACTCGGATC 432
QY 2152 TCTCTCTGAATCCACGACGCTATCCCTTTAGTTGAATAGGTTGAAGCTGTTC 2211
Db 433 TCTCTCTGAATCCACGACGCTATCCCTTTAGTTGAATAGGTTGAAGCTGTTC 492
QY 2212 AAGCCCAAGCAAAATGCACACCCCTTTTCTGTGGCAATCGTCTCTGATGTTGATCA 2271
Db 493 AAGCCCAAGCAAAATGCACACCCCTTTTCTGTGGCAATCGTCTCTGATGTTGATCA 552
QY 2272 TATTAGAAAGGGAAGATGTTAAGATATGTGGCTGTGGGTTACACAGGGTGCCTGCAGCG 2331
Db 553 TATTAGAAAGGGAAGATGTTAAGATATGTGGCTGTGGGTTACACAGGGTGCCTGCAGCG 612
QY 2332 GTAAATATTTTAGAAATATATATCAATACTCAACTCACTCACTCACTCACTCACTCACT 2391
Db 613 GTAAATATTTTAGAAATATATATCAATACTCAACTCACTCACTCACTCACTCACTCACT 672
QY 2392 ATTAATTTTTTTTTTTCTTTTAAAGAGAAAGCAGGCTTTTCTAGACTTTTAAAGAAATAAAGT 2451
Db 673 ATTAATTTTTTTTTTTCTTTTAAAGAGAAAGCAGGCTTTTCTAGACTTTTAAAGAAATAAAGT 732
QY 2452 CTTTGGGAGGCTCTCACGGTGTAGAGGAGCTTTTGGAGGCCACCGCCACAAA 2502
Db 733 CTTTGGGAGGCTCTCACGGTGTAGAGGAGCTTTTGGAGGCCACCGCCACAAA 783
```

```
RESULT 4
BX221568
LOCUS
DEFINITION
AGENCOURT_7559207 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6046477
5', mRNA sequence.
ACCESSION
BX221568
VERSION
BX221568.1 GI:20402968
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM13292 row: f column: 14
High quality sequence stop: 637.
Location/Qualifiers
1..851

FEATURES
source
1..851
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6046477"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 72"
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."

ORIGIN
Query Match 21.9%; Score 747; DB 13; Length 851;
Best Local Similarity 99.9%; Pred. No. 5.9e-285;
Matches 797; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 47 GCGGGCTTCGGGGAAGAGACGGATGATGAACAAGCTTTACATCGGAACCTGAGCCCGC 106
Db 1 GCGGGCTTCGGGGAAGAGACGGATGATGAACAAGCTTTACATCGGAACCTGAGCCCGC 60
QY 107 CCCTACCGCGACGACCTCGGCGAGCTCTTTGGGGACAGGAAGTGCCTTCGGGGAC 166
Db 61 CCCTACCGCGCGACCTCGGCGAGCTCTTTGGGGACAGGAAGTGCCTTCGGGGAC 120
QY 167 AGGTCTCTGCTGAAGTCCGGCTACGCTTCGTGGACTCCCGACACAGAACTGGCCATCC 226
Db 121 AGGTCTCTGCTGAAGTCCGGCTACGCTTCGTGGACTCCCGACACAGAACTGGCCATCC 180
QY 227 GCGGCATCGAGACCTCTCGGGTAAAGTGAATTGCATGGGAAATCATGGAAGTTGATT 286
Db 181 GCGGCATCGAGACCTCTCGGGTAAAGTGAATTGCATGGGAAATCATGGAAGTTGATT 240
QY 287 ACTCAGTCTTAAAGCTAAGAGCAGGAGGAGAAATTCAGATTCGAAACATCCCTCTCACC 346
Db 241 ACTCAGTCTTAAAGCTAAGAGCAGGAGGAGAAATTCAGATTCGAAACATCCCTCTCACC 300
QY 347 TGCAGTGGAGGTGTGATGCACTTTTGGCTCAATATGGACAGTGGAGATGTGGAAC 406
Db 301 TGCAGTGGAGGTGTGATGCACTTTTGGCTCAATATGGACAGTGGAGATGTGGAAC 360
QY 407 AAGTCAACACAGACACAGAAACCGCGTGTCAACGTGCATATGCAACACAGAGAAG 466
Db 361 AAGTCAACACAGACACAGAAACCGCGTGTCAACGTGCATATGCAACACAGAGAAG 420
QY 467 CAAAAATAGCCATGAGAGCTAAGCGGCGATCAGTTTGAACTACTCTCTCAAGTTT 526
Db 421 CAAAAATAGCCATGAGAGCTAAGCGGCGATCAGTTTGAACTACTCTCTCAAGTTT 480
QY 527 CCTACATCCCGGATGAAGAGTGAAGTCTCCCTTCGCCCTCAGCGAGCCAGCGTGGG 586
Db 481 CCTACATCCCGGATGAAGAGTGAAGTCTCCCTTCGCCCTCAGCGAGCCAGCGTGGG 540
QY 587 ACCACTCTTCCCGGAGCAAGGCCACGCCCCCTGGGGGCACTTCTCAGGCCACAGAGATTG 646

Db 541 ACCACTCTTCCCGGAGCAAGGCCACGCTCTGGGGGCACTTCTCAGGCCACAGAGATTG 600
QY 647 ATTTCCTCCGCTCGGATCCTGTCCTCCACCCAGTTTGTGGTCCCATCATCGGAAAGGAGG 706
Db 601 ATTTCCTCCGCTCGGATCCTGTCCTCCACCCAGTTTGTGGTCCCATCATCGGAAAGGAGG 660
QY 707 GCTTGACCATTAAGAACATCACTAAGCAGACCCAGTCCCGGCTAGATATCCATAGAAAG 766
Db 661 GCTTGACCATTAAGAACATCACTAAGCAGACCCAGTCCCGGCTAGATATCCATAGAAAG 720
QY 767 AGAACTCTGGAGCTGCAGAGAAAGCTGTCCATCATCCACCCAGGAGGAGCTTCTG 826
Db 721 AGAACTCTGGAGCTGCAGAGAAAGCTGTCCATCATCCATGCCACCCAGGAGGACTTCTG 780
QY 827 AAGCATGCCGATGATTC 844
Db 781 AAGCATGCCGATGATTC 798

RESULT 5
LOCUS CD243351 940 bp mRNA linear EST 22-MAY-2003
DEFINITION AGENCOURT 14121251 NIH MGC 180 Homo sapiens cDNA clone
IMAGE:30383283 5', mRNA sequence.
CD243351

ACCESSION CD243351.1 GI:31003815
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 940)
AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov

Tissue Procurement: Dr. Michael Brownstein
cDNA Library Preparation: Invitrogen Corp
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: NDAM446 row: o column: 04
High quality sequence start: 18
High quality sequence stop: 695.
Location/Qualifiers
1..940
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30383283"
/lab_host="DH10B-Ton A (T1 and T5 phage resistances)"
/clone_lib="NIH MGC 180"
/note="Organ: Testis; Vector: pCMV-SPORT6.1; Site 1: NotI;
Site 2: EcoRV (destroyed); Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.68 kb. Library was
constructed by (Invitrogen). Note: this is a NIH_MGC
Library."

FEATURES
source

ORIGIN

Query Match 21.5%; Score 734; DB 14; Length 940;
Best Local Similarity 100.0%; Pred. No. 7.7e-280;
Matches 734; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 125 TCGGGCAGCTCTTTGGGGACAGGAAGCTGCCCTTCGGGGACAGGTCCTCTGTAAGTCCG 184
Db 28 TCGGGCAGCTCTTTGGGGACAGGAAGCTGCCCTTCGGGGACAGGTCCTCTGTAAGTCCG 87
QY 185 GCTACGCTTCTGTGGACTACCCGACAGAACTGGGCGCATCCGCGCATCGAGACCTCT 244

Db 88 GCTACGCTTCGTGGACTACCCGACCAAGACTGGGCCATCCGGCCATCGAGACCTCT 147
QY 245 CGGTTAAAGTGGAAATTCATGCGAAATATCGAAGTTGATTACTCAGTCTCTAAAAAGC 304
Db 148 CGGTTAAAGTGGAAATTCATGCGAAATATCGAAGTTGATTACTCAGTCTCTAAAAAGC 207
QY 305 TAAGGAGCAGGAAATTCAGATTGCAATTCGAAACATCCCTCTCCTCAGTGGAGGTGTGG 364
Db 208 TAAGGAGCAGGAAATTCAGATTGCAAAACATCCCTCTCCTCAGTGGAGGTGTGG 267
QY 365 ATGGACTTTTGGCTCAATATGCGACAGTGGAGATGCGAACAGTCAACACAGACACAG 424
Db 268 ATGGACTTTTGGCTCAATATGCGACAGTGGAGATGCGAACAGTCAACACAGACACAG 327
QY 425 AAACCGCGTGTCAACGTCAATATCAACAGAGAGAAAGCAAAATAGCCATGGAGA 484
Db 328 AAACCGCGTGTCAACGTCAATATCAACAGAGAGAAAGCAAAATAGCCATGGAGA 387
QY 485 AGCTAAGCGGGCATCAGTTTGAGAACTACTCCTTCAGATTTCCTACATCCCGATGAAG 544
Db 388 AGCTAAGCGGGCATCAGTTTGAGAACTACTCCTTCAGATTTCCTACATCCCGATGAAG 447
QY 545 AGGTGAGTCCCTTCGCCCTCAGCGAGCCAGCGTGGGACCACTCTTCCCGGAGC 604
Db 448 AGGTGAGTCCCTTCGCCCTCAGCGAGCCAGCGTGGGACCACTCTTCCCGGAGC 507
QY 605 AAGGCCACGCCCTCGGGGCACTTCTCAGCCAGACAGATTGATTTCCCGCTCGGATCC 664
Db 508 AAGGCCACGCCCTCGGGGCACTTCTCAGCCAGACAGATTGATTTCCCGCTCGGATCC 567
QY 665 TGGTCCCCACCCAGTTGTGTGCTGCTATCCGAAAGAGGGCTTCACCATAAAGAA 724
Db 568 TGGTCCCCACCCAGTTGTGTGCTGCTATCCGAAAGAGGGCTTCACCATAAAGAA 627
QY 725 TCATTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAGAGAACTCTGGAGCTCAG 784
Db 628 TCATTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAGAGAACTCTGGAGCTCAG 687
QY 785 AGAAGCTGTACCATCCATGCCACCCAGAGGGGATTTCTGAAGCATGCCGATGATTC 844
Db 688 AGAAGCTGTACCATCCATGCCACCCAGAGGGGATTTCTGAAGCATGCCGATGATTC 747
QY 845 TTGAAATCATGCAG 858
Db 748 TTGAAATCATGCAG 761

RESULT 6
CF593505
LOCUS CF593505 824 bp mRNA linear EST 26-SEP-2003
DEFINITION AGENCOURT_15622249 NIH_MGC_147 Homo sapiens cDNA clone
IMAGE:30531076 5', mRNA sequence.
ACCESSION CF593505
VERSION CF593505.1 GI:36347108
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 824)
NTH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapob-remail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
and advice from Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
plate: NDAM621 row: e column: 05
High quality sequence stop: 682.
FEATURES
Location/Qualifiers
source
1..824
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30531076"
/tissue_type="Human Placenta"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC_147"
/notes="Organ: placenta; Vector: pBluescriptR; Site 1:
all-XhoI; Site 2: BamH; Oligo-dT primed using primer
5'-TTTTTTTTTTTTTNN-3', size-selected for average
insert size 2.3 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NHGRI/NHGRI, National Institutes of Health). Note: This is
a NIH_MGC library."

ORIGIN

Query Match 21.0%; Score 715; DB 14; Length 824;
Best Local Similarity 100.0%; Pred. No. 2.8e-272;
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCAGCGGAGGAGCGCGGTACCGGGCGGGGAGCGCGGCTCTCGGGG 60
Db 38 GGCAGCGGAGGAGCGCGGTACCGGGCGGGGAGCGCGGCTCTCGGGG 97
QY 61 AAGAGACGATGATGAACAAAGCTTTTACATCGGAAACCTGAGCCCGCGTCACGCCGAC 120
Db 98 AAGAGACGATGATGAACAAAGCTTTTACATCGGAAACCTGAGCCCGCGTCACGCCGAC 157
QY 121 GACCTCGGAGCTCTTTGGGAGCAGAGAGCTGCCCTTGGCGGGAGAGGTCCTGCTGAAG 180
Db 158 GACCTCGGAGCTCTTTGGGAGCAGAGAGCTGCCCTTGGCGGGAGAGGTCCTGCTGAAG 217
QY 181 TCCGGCTACGCTTCGTGGACTACCCGACCAAGCTGGGCCATCCGCGCCATCGAGACC 240
Db 218 TCCGGCTACGCTTCGTGGACTACCCGACCAAGCTGGGCCATCCGCGCCATCGAGACC 277
QY 241 CTCTCGGGTAAAGTGGAAATTCATGGAAATCATGGAAGTTGATTACTCAGTCTCTAAA 300
Db 278 CTCTCGGGTAAAGTGGAAATTCATGGAAATCATGGAAGTTGATTACTCAGTCTCTAAA 337
QY 301 AAGCTAAGGAGCGGAAATTCAGATTGGAACATCCCTCTCCTCAGCTGAGTGGAGGTG 360
Db 338 AAGCTAAGGAGCGGAAATTCAGATTGGAACATCCCTCTCCTCAGCTGAGTGGAGGTG 397
QY 361 TTGGATGACTTTTGGCTCAATATGGACAGTGGAGATGTGAAACAAGTCAACACAGAC 420
Db 398 TTGGATGACTTTTGGCTCAATATGGACAGTGGAGATGTGAAACAAGTCAACACAGAC 457
QY 421 ACAGAAACCGCGTTGTCAACGTCACATATGCAACAGAGAGAGAAAGCAAAATAGCCATG 480
Db 458 ACAGAAACCGCGTTGTCAACGTCACATATGCAACAGAGAGAGAAAGCAAAATAGCCATG 517
QY 481 GAGAAGCTAAGCGGGCATCAGTTTGAAGTACTCTTCAAGATTTCCTACATCCCGAT 540
Db 518 GAGAAGCTAAGCGGGCATCAGTTTGAAGTACTCTTCAAGATTTCCTACATCCCGAT 577
QY 541 GAGAGGTGAGTCCCTTCGCCCTCAGCGAGCCAGCGTGGGACCACTCTTCCCGG 600
Db 578 GAGAGGTGAGTCCCTTCGCCCTCAGCGAGCCAGCGTGGGACCACTCTTCCCGG 637
QY 601 GAGCAGGCCACGCCCTCGGGGCACTTCTCAGGCCAGACAGATTGATTTCCCGCTGCGG 660
Db 638 GAGCAGGCCACGCCCTCGGGGCACTTCTCAGGCCAGACAGATTGATTTCCCGCTGCGG 697

```
QY 661 ATCCGTGTCCTCCACCCAGTTTCTTGTGTCATCATCGGAAGAGGGCTTGACCA 715
|||||
Db 698 ATCCGTGTCCTCCACCCAGTTTCTTGTGTCATCATCGGAAGAGGGCTTGACCA 752
|||||

RESULT 7
BM806132 1089 bp mRNA linear EST 05-MAR-2002
LOCUS AGENCOURT_6553922 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:555652
DEFINITION 5', mRNA sequence.
ACCESSION BM806132
VERSION BM806132.1 GI:19122955
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1089)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12276 row: c column: 13
High quality sequence stop: 688.
FEATURES
Location/Qualifiers
1..1089
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:555652"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.1 kb."
ORIGIN
Query Match 20.6%; Score 702; DB 12; Length 1089;
Best Local Similarity 100.0%; Pred. No. 3.3e-267;
Matches 702; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 201 CTACCCCGACCAAGACTGGCCATCGGCCATCGAGACCCTCTCGGTAAGTGGAATT 260
|||||
Db 1 CTACCCCGACCAAGACTGGCCATCGGCCATCGAGACCCTCTCGGTAAGTGGAATT 60
|||||

QY 261 GCATGGGAAATCATGGAAGTTGATTACTCAGTCTCTAAAAGCTAAGGACAGGAAAT 320
|||||
Db 61 GCATGGGAAATCATGGAAGTTGATTACTCAGTCTCTAAAAGCTAAGGACAGGAAAT 120
|||||

QY 321 TCAGATTGGAACATCCCTCCTCAGTGGAGGTGTGATGACATTTTGGCTCA 380
|||||
Db 121 TCAGATTGGAACATCCCTCCTCAGTGGAGGTGTGATGACATTTTGGCTCA 180
|||||

QY 381 ATATGGGACAGTGGAGATTGGAACAAGTCAACACACACAGAAACCGCGTGTGCA 440
|||||
Db 181 ATATGGGACAGTGGAGATTGGAACAAGTCAACACACACAGAAACCGCGTGTGCA 240
|||||

QY 441 CGTCACATATGCAACAGAGAAAGAGAAAATGCCATGAGAGCTTAACGGGCATCA 500
|||||
Db 241 CGTCACATATGCAACAGAGAAAGAGAAAATGCCATGAGAGCTTAACGGGCATCA 300
|||||

QY 501 GTTTGAGAACTACTCTTCAGATTTCCTACATCCCGATCAAGAGGTGAGTCCCTTC 560
|||||
```

```
Db 301 GTTTGAGAACTACTCTTCAGATTTCCTACATCCCGATGAAGAGTGAGTCCCTTC 360
QY 561 GCCCCTCAGGAGCCCGAGCGTGGGACCACTTTCCTCCGGGAGGAGCCAGCCCTTGG 620
|||||
Db 361 GCCCCTCAGGAGCCCGAGCGTGGGACCACTTTCCTCCGGGAGGAGCCAGCCCTTGG 420
|||||

QY 621 GGGCACTTCTCAGGCGACACAGATTGATTTCCGCTCGGATCCCTGTTCCCAACCA 580
|||||
Db 421 GGGCACTTCTCAGGCGACACAGATTGATTTCCGCTCGGATCCCTGTTCCCAACCA 480
|||||

QY 681 TGTGTTGCCATCATCGGAAAGAGGGCTTGACCATAAAGAACATCACTAAGACAGACCA 740
|||||
Db 481 TGTGTTGCCATCATCGGAAAGAGGGCTTGACCATAAAGAACATCACTAAGACAGACCA 540
|||||

QY 741 GTCCGGGTAGATATCCATAGAAAGAGAACTCTGGAGCTGCAGAGAGCCTGTCCAT 800
|||||
Db 541 GTCCGGGTAGATATCCATAGAAAGAGAACTCTGGAGCTGCAGAGAGCCTGTCCAT 600
|||||

QY 801 CCATGCCACCCCGAGAGGGGACTTCTGAAGCATGCCGATGATTTCTTGAATCATGCAGAA 860
|||||
Db 601 CCATGCCACCCCGAGAGGGGACTTCTGAAGCATGCCGATGATTTCTTGAATCATGCAGAA 660
|||||

QY 861 AGAGCGAGATGAGACCAAACTAGCCGAGAGAGATTCTCTGAA 902
|||||
Db 661 AGAGCGAGATGAGACCAAACTAGCCGAGAGAGATTCTCTGAA 702
|||||

RESULT 8
BM561242
LOCUS AGENCOURT_6567353 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:5739550
DEFINITION 5', mRNA sequence.
ACCESSION BM561242
VERSION BM561242.1 GI:18906337
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 999)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12753 row: a column: 23
High quality sequence stop: 679.
FEATURES
Location/Qualifiers
1..999
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5739550"
/tissue_type="duodenal adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_88"
/note="Organ: small intestine; Vector: pCMV-SPORT6;
Site 1: NotI; Site 2: SalI; Cloned unidirectionally;
oligo-dr primed. Average insert size 1.767 Kb. Library
enriched for full-length clones and constructed by Life
Technologies. Note: this is a NIH_MGC Library."
ORIGIN
Query Match 20.5%; Score 698; DB 12; Length 999;
Best Local Similarity 100.0%; Pred. No. 1.3e-265;
Matches 698; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```


FEATURES source POLYA=Yes Location/Qualifiers
1. .647 /organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-DU0-aac-j-06-0-UI"
/tissue_type="Primary Lung Epithelial Cells"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-CF-DU0"
/notes="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; UI-CF-DU0 is a cDNA library containing the following tissue(s): Primary lung Epithelial Cells The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GGCTGTAGGC.
TAG_TISSUE=Lung Epithelial Cells Tissue nos 359-368
TAG_LIB=UI-CF-DU0
TAG_SEQ=GGCTGTAGGC"

ORIGIN
Query Match 18.6%; Score 634; DB 13; Length 647;
Best Local Similarity 100.0%; Pred. No. 3.8e-240;
Matches 634; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 371 TTTTGGCTCAATATGGACAGTGGAGATGTGGAACAAGTCAACACAGACAGACAAACCG 430
DB 14 TTTTGGCTCAATATGGACAGTGGAGATGTGGAACAAGTCAACACAGACAGACAAACCG 73
QY 431 CCGTTGTCAACGTCACATATGCAACAGAGAAGCAAAATAGCCATCGGAAGCTAA 490
DB 74 CCGTTGTCAACGTCACATATGCAACAGAGAAGCAAAATAGCCATCGGAAGCTAA 133
QY 491 GCGGGCATAGTTTGAGAACTACTCTTCAAGATTTCTACATCCCGGATCGAGAGTGA 550
DB 134 GCGGGCATAGTTTGAGAACTACTCTTCAAGATTTCTACATCCCGGATCGAGAGTGA 193
QY 551 GCTCCCTTCGCCCTTCAGCGAGCCAGCGTGGGACCACTCTTCCCGGAGCAAGGCC 610
DB 194 GCTCCCTTCGCCCTTCAGCGAGCCAGCGTGGGACCACTCTTCCCGGAGCAAGGCC 253
QY 611 ACGCCCTGGGGCACTTCTCAGGCGACAGAGATTGATTTCCCGCTCGGATCTGGTCC 670
DB 254 ACGCCCTGGGGCACTTCTCAGGCGACAGAGATTGATTTCCCGCTCGGATCTGGTCC 313
QY 671 CCACCCAGTTTGTGTGGCCATCATCGGAAGAGGGCTTGACCATAAAGACATCACTA 730
DB 314 CCACCCAGTTTGTGTGGCCATCATCGGAAGAGGGCTTGACCATAAAGACATCACTA 373
QY 731 AGCAGACCCAGTCCCGGTAGATATCCATAGAAAGAGAACTCTGGAGCTGCGAGAGC 790
DB 374 AGCAGACCCAGTCCCGGTAGATATCCATAGAAAGAGAACTCTGGAGCTGCGAGAGC 433
QY 791 CTGTACATCCATGCCACCCGAGGGGACTTCTGAAGCATGCCGATGATTTGAAA 850
DB 434 CTGTACATCCATGCCACCCGAGGGGACTTCTGAAGCATGCCGATGATTTGAAA 493
QY 851 TCATGAGAGAGGCGAGATGAGCAAACTAGCCGAGAGATTCCTCTGAATCTTGG 910
DB 494 TCATGAGAGAGGCGAGATGAGCAAACTAGCCGAGAGATTCCTCTGAATCTTGG 553
QY 911 CACACATGGCTTGGTGGAGACTGATTGGAAAGAGGAGAGAAATTTGAAGAAATTG 970
DB 554 CACACATGGCTTGGTGGAGACTGATTGGAAAGAGGAGAGAAATTTGAAGAAATTG 613

QY 971 AACATGAACAGGAGCAACAGATAACAATCTCATC 1004
DB 614 AACATGAACAGGAGCAACAGATAACAATCTCATC 647

RESULT 12
BG748346 889 bp mRNA linear EST 15-MAY-2001
LOCUS 602705902F1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:4842553 5',
DEFINITION mRNA sequence.
ACCESSION BG748346
VERSION BG748346.1 GI:114058999
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 889)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCMI677 row: c column: 02
High quality sequence stop: 773.

FEATURES source
1. .889 Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4842553"
/tissue_type="normal pigmented retinal epithelium"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_43"
/note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library. |"

ORIGIN
Query Match 18.1%; Score 617; DB 12; Length 889;
Best Local Similarity 99.7%; Pred. No. 1.7e-233;
Matches 717; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 72 GATGAACAAGCTTTACATCGGGAACCTGAGCCCGCGTCACCGCAGACCTCCGGCA 131
DB 2 GATGAACAAGCTTTACATCGGGAACCTGAGCCCGCGTCACCGCAGACCTCCGGCA 61
QY 132 GCTCTTTGGGACAGAGAGCTGCCCTGGGGACAGAGTCTCTGTAAGTCCGGCTACGC 191
DB 62 GCTCTTTGGGACAGAGAGCTGCCCTGGGGACAGAGTCTCTGTAAGTCCGGCTACGC 121
QY 192 CTTCTGTGGACTACCCCGACAGAACTGGCGCATCGCGCATCGAGACCTCTCGGGTAA 251
DB 122 CTTCTGTGGACTACCCCGACAGAACTGGCGCATCGCGCATCGAGACCTCTCGGGTAA 181
QY 252 AGTGAATTCGATGGGAAATCATGGAAGTTGATTACTCAGTCTCTAAAGAGTAAGGAG 311
DB 182 AGTGAATTCGATGGGAAATCATGGAAGTTGATTACTCAGTCTCTAAAGAGTAAGGAG 241
QY 312 CAGGAAATTCAGATTCGGAACAATCCCTCTCTACCTGCAGTGGGAGGTGTGGATGACT 371

Db 242 CAGGAAATTCAGATTCGAAACATCCCTCCTCACCTGCAGTGGGAGGTGTTGGATGACT 301
Qy 372 TTTGGCTCAATATGGGACAGTGGAGAAATGTGGAACTCAACAGACACAGAAACCGC 431
Db 302 TTTGGCTCAATATGGGACAGTGGAGAAATGTGGAACTCAACAGACACAGAAACCGC 361
Qy 432 CGTTGTCACAGTCACATATGCAACAAAGAGAGAAACAAAATAGCCATGGAGAACTAAG 491
Db 362 CGTTGTCACAGTCACATATGCAACAAAGAGAGAAACAAAATAGCCATGGAGAACTAAG 421
Qy 492 CGGGCATCAGTTTGAGAACTACTCCTTCAAGATTTCTTACATCCCGGATGAAGAGGTGAG 551
Db 422 CGGGCATCAGTTTGAGAACTACTCCTTCAAGATTTCTTACATCCCGGATGAAGAGGTGAG 481
Qy 552 CTCCTCTCGCCCTCAGCGAGCCAGCGTGGGACCACTCTTCCCGGGAGCAAGGCCA 611
Db 482 CTCCTCTCGCCCTCAGCGAGCCAGCGTGGGACCACTCTTCCCGGGAGCAAGGCCA 541
Qy 612 CGCCCTCGGGGCACCTTCTCAGGCGACAGACAGATTGATTTCCCGCTCGCGATCCTGGTCCC 671
Db 542 CGCCCTCGGGGCACCTTCTCAGGCGACAGACAGATTGATTTCCCGCTCGCGATCCTGGTCCC 601
Qy 672 CACCCAGTTTGTGTGTCATCATCGGAAAGAGAGGCTTGACCAATAAAGAAACATCACTAA 731
Db 602 CAACAGTTTGTGTGTCATCATCGGAAAGAGAGGCTTGACCAATAAAGAAACATCACTAA 661
Qy 732 GCAGACCCAGTCCCGGGTAGATATCCATAGAAAGAGAACTCTGGAGCTGCAGAGAAC 790
Db 662 GCAGACCCAGTCCCGGGTAGATATCCATAGAAAGAGAACTCTGGAGCTGCAGAGAAC 720

RESULT 13
BG575889
LOCUS 602598315F1 NIH_MGC_87 Homo sapiens cdna clone IMAGE:4707243 5',
DEFINITION mRNA sequence.
ACCESSION BG575889.1 GI:13583542
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 781)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
cdna Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10576 row: m column: 04
High quality sequence stop: 773.
Location/Qualifiers
1..781
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4707243"
/tissue_type="mammary adenocarcinoma, cell line"
/lab_host="DHI0B (phage-resistant)"
/clone_lib="NIH_MGC_87"
/note="Organ: breast; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.383 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

FEATURES
source

ORIGIN
Query Match 17.4%; Score 594; DB 12; Length 781;
Best Local Similarity 99.7%; Pred. No. 2.3e-224;
Matches 694; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1845 AGTCGCTTCACAGCGCAGCAAGTGGAGTCCACAGGCAACAGCAAAACAAACCGGATGAAT 1904
Db 25 AGTCGCTTCACAGCGCAGCAAGTGGAGTCCACAGGCAACAGCAAAACAAACCGGATGAAT 84
Qy 1905 GTAGCCCTTCCAAACACCTGACAGAAATGAGACAAACGAGCCAGCAGATCGGGAGCAAA 1964
Db 85 GTAGTCCTTCCAAACACCTGACAGAAATGAGACAAACGAGCCAGCAGATCGGGAGCAAA 144
Qy 1965 CCAAGACCATCTGAGGATCAGAAAGTCTCGGAGGGGGCCAGGGACTCTGCCAGGGGCC 2024
Db 145 CCAAGACCATCTGAGGATCAGAAAGTCTCGGAGGGGGCCAGGGACTCTGCCAGGGGCC 204
Qy 2025 TGAGAACCCCAAGGGGCCGAGGAGGGGGGGAAGGTACGCCAGGTTTCCCAAGAACCAACG 2084
Db 205 TGAGAACCCCAAGGGGCCGAGGAGGGGGGGAAGGTACGCCAGGTTTCCCAAGAACCAACG 264
Qy 2085 AGCCCGGCTCTCCGCCCCCAGGGCTTCTGCAAGCTTCAGGCATCCACTTCCACCATCCAC 2144
Db 265 AGCCCGGCTCTCCGCCCCCAGGGCTTCTGCAAGCTTCAGGCATCCACTTCCACCATCCAC 324
Qy 2145 TCGGATCTCTCCTGAACTCCCAAGGCTATCCCTTTTCTGTCGCAAAATCGTCTCTGTACATGT 2204
Db 325 TCGGATCTCTCCTGAACTCCCAAGGCTATCCCTTTTCTGTCGCAAAATCGTCTCTGTACATGT 384
Qy 2205 GTGTTCAAAGCCAAAGCAAAATGCACACCCCTTTTCTGTGGCAAAATCGTCTCTGTACATGT 2264
Db 385 GTGTTCAAAGCCAAAGCAAAATGCACACCCCTTTTCTGTGGCAAAATCGTCTCTGTACATGT 444
Qy 2265 GTGTACATATTAGAAAGGAAAGATGTTAAGATATGTGGCCCTGTGGTTPACACAGGGTGCC 2324
Db 445 GTGTACATATTAGAAAGGAAAGATGTTAAGATATGTGGCCCTGTGGTTPACACAGGGTGCC 504
Qy 2325 TGCAGCGGTAAATATATTAGAAATAATATCAAAATCAAACTCAACTTCACTTCAATTTTAA 2384
Db 505 TGCAGCGGTAAATATATTAGAAATAATATCAAAATCAAACTCAACTTCACTTCAATTTTAA 564
Qy 2385 ATCAATATTAAATTTTTTTTCTTTTAAAGAGAAAGCAGGCTTTTCTAGACTTTTAAAGA 2444
Db 565 ATCAATATTAAATTTTTTTTCTTTTAAAGAGAAAGCAGGCTTTTCTAGACTTTTAAAGA 624
Qy 2445 ATAAAGTCTTTGGAGGTCTCACCGTGTAGAGAGAGCTTTTGAGGCCACCGGCACAAAT 2504
Db 625 ATACAGTCTTTGGAGGTCTCACCGTGTAGAGAGAGCTTTTGAGGCCACCGGCACAAAT 684
Qy 2505 TCACCCAGAGGGAATTCGTGCGAAGGACACTCAC 2540
Db 685 TCACCCAGAGGGAATTCGTGCGAAGGACACTCAC 720
RESULT 14
LOCUS BUI56245
DEFINITION BUI56245.1 GI:22669777
ACCESSION BUI56245
VERSION BUI56245.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov

Qy	2500	AAAATTCA	CCCAAGG	AGAAATCT	CGTCGGA	AGGACACT	CA	CGGAG	AGTTCT	GGATCA	CCCTG	2559	
Db	421	AAAATTTC	CAAGGAG	AAATCT	CGTCGGA	AGGACACT	CA	CGGAG	AGTTCT	GGATCA	CCCTG	480	
Qy	2560	TGTA	TGTCAC	CAGAAG	GGATAC	CGTCTCC	TTG	AGAG	AGAACTCT	GTCACT	CTCTCAT	GCC	2619
Db	481	TGTA	TGTCAC	CAGAAG	GGATAC	CGTCTCC	TTG	AGAG	AGAACTCT	GTCACT	CTCTCAT	GCC	540
Qy	2620	TGTC	TAGTCAT	ACACCA	CTATTC	TCTTT	TGCTT	CAC	CAGGTTTAA	ACTGG	TTTTTGC	AAT	2679
Db	541	TGTC	TAGTCAT	ACACCA	CTATTC	TCTTT	TGCTT	CAC	CAGGTTTAA	ACTGG	TTTTTGC	AAT	600
Qy	2680	CTGCTA	TATAA	ATTC	TCTGT	CTCTCT	CTCTGT	TTATCTC				2715	
Db	601	CTGCTA	TATAA	ATTC	TCTGT	CTCTCT	CTCTGT	TTATCTC				636	

Search completed: July 21, 2004, 19:23:44
Job time : 8238 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: July 16, 2004, 10:30:58 ; Search time 149 Seconds
(without alignments)
12940.296 Million cell updates/sec

Title: US-09-270-437D-6
Perfect score: 6121
Sequence: 1 GGCAGCGAGGAGCGAGGA.....AACCTGAAATGTTTATT 3412

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 3172214

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2.1/USPTO.spool.p/US09270437/runat.16072004.113048.13308/app.query.fasta.1.3591
-DB=A Geneseq 29Jan04 -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS-human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09270437@cgn.1.253@runat.16072004.113048.13308 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A Geneseq 29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3094	50.5	614	4 ABG06794	Abg06794 Novel hum
2	3065.5	50.1	620	4 AAU16163	Aau16163 Human nov
3	3065.5	50.1	620	6 ABUS5232	Abu5232 Human nov
4	2917.5	47.7	594	4 ABG06795	Abg06795 Novel hum
5	2806.5	45.9	556	5 ABG96346	Abg96346 Human ova
6	2553	41.7	555	6 ABG89799	Abg89799 Novel hum
7	2019.5	33.0	577	2 AAY30649	Aay30649 A murine
8	1957.5	32.0	579	3 AAB11365	Aab11365 Human lun
9	1957.5	32.0	579	5 ABB75053	Abb75053 Human lun
10	1957.5	32.0	579	5 ABB74997	Abb74997 Human lun

11	1957.5	32.0	579	5 ABB75054	Abb75054 Human lun
12	1957.5	32.0	579	5 ABP61917	Abp61917 Human lun
13	1957.5	32.0	579	5 ABP61974	Abp61974 Human lun
14	1957.5	32.0	579	5 ABP61973	Abp61973 Human lun
15	1957.5	32.0	579	7 ADA28536	Ada28536 Recombina
16	1957.5	32.0	579	7 ADA28539	Ada28539 Recombina
17	1957.5	32.0	579	7 ADA28438	Ada28438 Human lun
18	1952.5	31.9	586	5 ABB75048	Abb75048 Human lun
19	1952.5	31.9	586	5 ABP61968	Abp61968 Human lun
20	1952.5	31.9	586	7 ADA28517	Ada28517 Recombina
21	1950.5	31.9	579	3 AAB11328	Aab11328 Human lun
22	1950.5	31.9	579	5 ABB74960	Abb74960 Human lun
23	1950.5	31.9	579	5 ABP61880	Abp61880 Human lun
24	1950.5	31.9	579	7 ADA28286	Ada28286 Human lun
25	1950.5	31.9	579	7 ADP14066	Adp14066 Human src
26	1950.5	31.9	579	7 ADE53471	Ades3471 Human lun
27	1877.5	30.7	619	4 ABG21963	Abg21963 Novel hum
28	1804.5	29.5	583	4 ABG12592	Abg12592 Novel hum
29	1161	19.0	319	4 AAM93826	Aam93826 Human pol
30	910	14.9	192	4 AAU16164	Aau16164 Human nov
31	910	14.9	192	6 ABUS5233	Abu5233 Human nov
32	898	14.7	261	4 AAU16161	Aau16161 Human nov
33	898	14.7	261	6 ABUS5230	Abu5230 Human nov
34	868	14.2	250	4 AAU16579	Aau16579 Human nov
35	868	14.2	250	6 ABUS5648	Abu5648 Human nov
36	865.5	14.1	558	4 ABBS8367	Abbs8367 Drosophil
37	840.5	13.7	187	6 AAO23971	Aao23971 Human IGF
38	803	13.1	171	4 AAU16166	Aau16166 Human nov
39	803	13.1	171	6 ABUS5235	Abu5235 Human nov
40	750	12.3	171	4 AAU16583	Aau16583 Human nov
41	750	12.3	171	6 ABUS5652	Abu5652 Human nov
42	501	8.2	97	3 AAG03261	Aag03261 Human sec
43	458	7.5	148	4 ABG21962	Abg21962 Novel hum
44	452	7.4	266	4 ABG12593	Abg12593 Novel hum
45	448	7.3	209	4 ABG21961	Abg21961 Novel hum

ALIGNMENTS

RESULT 1
ABG06794
ID ABG06794 standard; protein; 614 AA.
AC ABG06794;
XX XX
DT 13-FEB-2002 (first entry)
XX XX
DE Novel human diagnostic protein #6785.
XX XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX XX
OS Homo sapiens.
XX XX
PN WC200175067-A2.
XX XX
PD 11-OCT-2001.
XX XX
PF 30-MAR-2001; 2001WO-US008631.
XX XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX XX
(HYSE-) HYSEQ INC.
PA Drmanac RT, Liu C, Tang YT;
PI
XX
XX WPI; 2001-639362/73.
DR N-PSDB; AAS70981.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess

PT biodiversity.
 PS Claim 20; SEQ ID NO 37153; 103pp; English.
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic amino acid sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 614 AA;

Alignment Scores:
 Pred. No.: 5.53e-271 Length: 614
 Score: 3094.00 Matches: 610
 Percent Similarity: 99.03% Conservative: 0
 Best Local Similarity: 99.03% Mismatches: 4
 Query Match: 50.55% Indels: 2
 DB: 4 Gaps: 0

US-09-270-437d-6 (1-3412) x ABG06794 (1-614)

QY 23 GCCGGGTACCGGGCGGGGAGCGCGGCTCTCGGGAGAGACGGATGATGAACAGC 82
 Db 1 AlaGlyTyrArgArg-ArgGluProArgPro-GlnGlyLysArgArgMetMetAsnLysL 20

QY 83 TTTACATCGGAACCTGAGCCCGCGTACCGCGAGCACCTCCGGCAGCTCTTTGGGG 142
 Db 20 euTyrIleGlyAsnLeuSerProAlaValThrAlaAspLeuArgGlnLeuPheGlyA 40

QY 143 ACAGGAGCTCCCTCGGGGAGAGCTCTGCTGAAGTCCGGCTACGCTTCGTGACT 202
 Db 40 spArgLysLeuProLeuAlaGlyGlnValLeuLysSerGlyTyrAlaPheValAspT 60

QY 203 ACCCGGACCAAGTGGGCCATCCGGCCATCGAGACCTCTCGGTAAAGTGAATTGC 262
 Db 60 YrProAspGlnAsnTrpAlaIleArgAlaIleGluThrLeuSerGlyLysValGluLeuH 80

QY 263 ATGGGAAATCATGGAATGATTACTCAGTCTCTAAAGCTAAGGACGAGGAAATTC 322
 Db 80 isGlyLysIleMetGluValAspTyrSerValSerLysLysLeuArgSerArgLysIleG 100

QY 323 AGATTGGAACATCCCTCCTCACCTGAGTGGAGGTGTGGATGGACTTTGGCTCAAT 382
 Db 100 InIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspGlyLeuLeuAlaGlnT 120

QY 383 ATGGGACAGTGGAGATGTGAAACAAGTCAACACAGACACAGAACCGCGCTTGTCAAG 442
 Db 120 YrGlyThrValGluAsnValGluGlnValAsnThrAspThrGluThrAlaValValAsnV 140

QY 443 TCACATATGCACACAGAGAGAGCAAAATAGCCATGGAGACCTTAAGCGGGCATCAGT 502
 Db 140 alThrTyrAlaThrArgGluGluAlaLysIleAlaMetGluLysLeuSerGlyHisGlnP 160

QY 503 TTGAGAACTACTCTTCAAGATTTCTTACATCCCGGATGAAGAGTGAAGCTCCCTTCGC 562
 Db 160 heGluAsnTyrSerPheLysIleSerTyrIleProAspGluGluValSerSerProSerP 180

QY 563 CCCCTCAGCAGCCAGCGTGGGAGCCACTCTTCCCGGAGACAGGCCACGCCCTGGG 622
 Db 180 roProGlnArgAlaGlnArgGlyAspHisSerSerArgGlnGlnGlyHisAlaProGlyG 200

QY 623 GCATCTTCCAGCCAGACAGATTGATTTCCCGCTGGGATCTGTCTCCCAACCCAGTTTG 682
 Db 200 lyThrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeuValProThrGlnPheV 220

QY 683 TTGTTGCCATCATCGAAAGAGGGCTTGACCATAAAGAACATCACATAAGCAGACCCAGT 742
 Db 220 alGlyAlaIleIleGlyLysGluGlyLeuThrIleLysAsnIleThrLysGlnThrGlnS 240

QY 743 CCCGGTAGATATCCATAGAAACAGAACTCTGAGCTGCAGAGAGCCTCTCCACCAACC 802
 Db 240 erArgValAspIleHisArgLysGluAsnSerGlyAlaAlaGluLysProValThrIleH 260

QY 803 ATGCACCCAGAGGGGACTTCTGAAGCATCCCGCATGATTCTTGAATCATGAGAGAAG 862
 Db 260 isAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeuGluIleMetGlnLysG 280

QY 863 AGCAGATGAGACCAACTAGCCGAGAGATTCTCTGAAATCTTGGCACACATGCTGCT 922
 Db 280 luAlaAspGluThrLysLeuAlaGluGluIleProLeuLysIleLeuAlaHisAsnGlyL 300

QY 923 TGGTTGGAAGACTGATTGGAAAGAGGACAGAAATTTGAAGAAATTTGAACATGAACAG 982
 Db 300 euValGlyArgLeuIleGlyLysGluArgAsnLeuLysIleGluHisGluThrG 320

QY 983 GGACCAAGATAACAATCTCATCTTTGAGGATTTGAGCATATACAACCCGGAAGAACCA 1042
 Db 320 lyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsnProGluArgThrI 340

QY 1043 TCATCTGAGGGCAGACAGTGTGAGCCCTGTGCGCAGTGTGAGATGAGATATGAAGAAGC 1102
 Db 340 leThrValLysGlyThrValGluAlaCysAlaSerAlaGluIleGluIleMetLysLysL 360

QY 1103 TGCCTGAGGCTTTGAAATGATGCTGGCTGTAAACCAACAGCAATCTGATCCAG 1162
 Db 360 euArgGluAlaPheGluAsnAspMetLeuAlaValAsnGlnGlnAlaAsnLeuIleProG 380

QY 1163 GGTGAAACCTCAGCGCATTGGCATCTTTTCAACAGGACTGTCGCTGCTATCTCCACAG 1222
 Db 380 lyLeuAsnLeuSerAlaLeuGlyIlePheSerThrGlyLeuSerValLeuSerProProA 400

QY 1223 CAGGGCCCGGGAGCTCCCGCTGCCCTACCGACCCCTTCTACCTACCCACCTCCGGAT 1282
 Db 400 laGlyProArgGlyAlaProProAlaAlaProTyrHisProPheThrThrHisSerGlyT 420

QY 1283 ACTTCTCAGCCTGTACCCCATCACAGTTTGGCCCGTTCGCCCATCATCCTCTTATC 1342
 Db 420 yrPheSerSerLeuTyrProHisGlnPheGlyProPheProHisHisSerTyrP 440

QY 1343 CAGACGAGGATGTGAATCTTTCATCCCAACCCAGGCTGTGGCGCCATCATCGGGA 1402
 Db 440 roGluGlnGluIleValAsnLeuPheIleProThrGlnAlaValGlyAlaIleIleGlyL 460

QY 1403 AGAAGGGGCACACATCAACAGCTGGGAGATTGGCGGAGCCTCTCATAGATTGCC 1462
 Db 460 yLysGlyAlaHisIleLysGlnLeuAlaArgPheAlaGlyAlaSerIleLysIleAlaP 480

QY 1463 CTGGGGAAGGCCACAGCTCAGCAAGGATGGTCATCATCCCGGCCACCCGGAAGCCC 1522
 Db 480 roAlaGluGlyProAspValSerGluArgMetValIleIleThrGlyProProGluAlaG 500

QY 1523 AGTTCAAGCCCGAGGACCGGATCTTTGGGAACTGGAAGAGGAACTCTTTTAAACCCCA 1582
 Db 500 lnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluGluAsnPhePheAsnProL 520

QY 1583 AAGAAGAGTGAAGCTGGAAGCGCATATCAGAGTGCCTCTTCCACAGCTGGCCGGTGA 1642
 Db 520 ysGluGluValLysLeuGluAlaHisIleArgValProSerSerThrAlaGlyArgValI 540

QY	1643	TTGGCAAGTGGCAAGACGGTGAACGAACTGACAGAACTTAACCACTGCAGAGTCAATCG	1702	PR	14-AUG-2000;	2000US-0225759P.
Db	540	leclYlysGlyGlyThrValAsnGluLeuGlnAsnLeuThrSerAlaGluValIleV	560	PR	18-AUG-2000;	2000US-0226279P.
QY	1703	TGCTCTGTCACCAAGCCAGATGAAAATGAGGAAGTGTATCGTTCAGAAATATCGGGCACT	1762	PR	22-AUG-2000;	2000US-0226681P.
Db	560	alProArgaspGlnThrProAspGluAsnGluGluValIleValArgIleIleGlyHisP	580	PR	22-AUG-2000;	2000US-0226868P.
QY	1763	TCATTGCTAGCCAGACTGCACAGCGAAGATCAGGAAATGTACAAACAGGTGAAGCAGC	1822	PR	23-AUG-2000;	2000US-0227182P.
Db	580	heheAlaSerGlnThrAlaGlnArgLysIleArgGluIleValGlnGlnValLysGlnG	600	PR	23-AUG-2000;	2000US-0227182P.
QY	1823	AGGACAGAAATACCTCAGGGAGTGCCTTCACAGCGCAGCAAG	1866	PR	30-AUG-2000;	2000US-0228924P.
Db	600	InGluGlnIlytyrProGlnGlyValAlaSerGlnArgSerLys	614	PR	01-SEP-2000;	2000US-0229287P.
RESULT 2				PR	01-SEP-2000;	2000US-0229343P.
AAU16163				PR	01-SEP-2000;	2000US-0229344P.
ID	AAU16163	standard; protein; 620 AA.		PR	01-SEP-2000;	2000US-0229345P.
XX	AC	AAU16163;		PR	05-SEP-2000;	2000US-0229509P.
XX	DT	07-NOV-2001 (first entry)		PR	05-SEP-2000;	2000US-0229513P.
XX	DE	Human novel secreted protein, Seq ID 1116.		PR	06-SEP-2000;	2000US-0230437P.
XX	XX	Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;		PR	06-SEP-2000;	2000US-0230438P.
KW	KW	cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;		PR	08-SEP-2000;	2000US-0231242P.
KW	KW	antibacterial; virucide; fungicide; ophthalmological; vulnerary;		PR	08-SEP-2000;	2000US-0231243P.
KW	KW	secreted protein; rheumatoid arthritis; hyperproliferative disorder;		PR	08-SEP-2000;	2000US-0231244P.
KW	KW	cardiovascular disorder; cardiac arrest; cerebrovascular disorder;		PR	08-SEP-2000;	2000US-0231413P.
KW	KW	cerebral ischaemia; angiogenesis; nervous system disorder;		PR	08-SEP-2000;	2000US-0231414P.
KW	KW	Alzheimer's disease; infection; ocular disorder; corneal infection;		PR	08-SEP-2000;	2000US-0232080P.
KW	KW	wound healing; epithelial cell proliferation; skin ageing; food additive;		PR	08-SEP-2000;	2000US-0232081P.
XX	XX	preservative; antiproliferative.		PR	14-SEP-2000;	2000US-0232397P.
XX	OS	Homo sapiens.		PR	14-SEP-2000;	2000US-0232398P.
XX	XX			PR	14-SEP-2000;	2000US-0232399P.
PN	WO200155322-A2.			PR	14-SEP-2000;	2000US-0232400P.
PD	02-AUG-2001.			PR	14-SEP-2000;	2000US-0232401P.
XX	17-JAN-2001; 2001WO-US001341.			PR	14-SEP-2000;	2000US-0233063P.
XX	31-JAN-2000; 2000US-0179065P.			PR	14-SEP-2000;	2000US-0233064P.
PR	04-FEB-2000; 2000US-0180628P.			PR	14-SEP-2000;	2000US-0233065P.
PR	24-FEB-2000; 2000US-0184664P.			PR	21-SEP-2000;	2000US-0234223P.
PR	02-MAR-2000; 2000US-0186350P.			PR	21-SEP-2000;	2000US-0234274P.
PR	16-MAR-2000; 2000US-0189874P.			PR	25-SEP-2000;	2000US-0234997P.
PR	17-MAR-2000; 2000US-0190076P.			PR	25-SEP-2000;	2000US-0234998P.
PR	18-APR-2000; 2000US-0198123P.			PR	26-SEP-2000;	2000US-0235484P.
PR	19-MAY-2000; 2000US-0205515P.			PR	26-SEP-2000;	2000US-0235484P.
PR	07-JUN-2000; 2000US-0209467P.			PR	27-SEP-2000;	2000US-0235834P.
PR	28-JUN-2000; 2000US-0214886P.			PR	27-SEP-2000;	2000US-0235836P.
PR	30-JUN-2000; 2000US-0215135P.			PR	29-SEP-2000;	2000US-0236327P.
PR	07-JUL-2000; 2000US-0216647P.			PR	29-SEP-2000;	2000US-0236368P.
PR	11-JUL-2000; 2000US-0217487P.			PR	29-SEP-2000;	2000US-0236369P.
PR	11-JUL-2000; 2000US-0217496P.			PR	29-SEP-2000;	2000US-0236370P.
PR	14-JUL-2000; 2000US-0218290P.			PR	02-OCT-2000;	2000US-0236802P.
PR	26-JUL-2000; 2000US-0220963P.			PR	02-OCT-2000;	2000US-0237037P.
PR	26-JUL-2000; 2000US-0220964P.			PR	02-OCT-2000;	2000US-0237038P.
PR	14-AUG-2000; 2000US-0224518P.			PR	02-OCT-2000;	2000US-0237039P.
PR	14-AUG-2000; 2000US-0224519P.			PR	02-OCT-2000;	2000US-0237040P.
PR	14-AUG-2000; 2000US-0225213P.			PR	13-OCT-2000;	2000US-0239355P.
PR	14-AUG-2000; 2000US-0225214P.			PR	13-OCT-2000;	2000US-0239377P.
PR	14-AUG-2000; 2000US-0225266P.			PR	20-OCT-2000;	2000US-0241808P.
PR	14-AUG-2000; 2000US-0225267P.			PR	20-OCT-2000;	2000US-0241809P.
PR	14-AUG-2000; 2000US-0225268P.			PR	20-OCT-2000;	2000US-0241826P.
PR	14-AUG-2000; 2000US-0225270P.			PR	01-NOV-2000;	2000US-0244617P.
PR	14-AUG-2000; 2000US-0225447P.			PR	08-NOV-2000;	2000US-0246474P.
PR	14-AUG-2000; 2000US-0225757P.			PR	08-NOV-2000;	2000US-0246475P.
PR	14-AUG-2000; 2000US-0225758P.			PR	08-NOV-2000;	2000US-0246476P.
PR	14-AUG-2000; 2000US-0225759P.			PR	08-NOV-2000;	2000US-0246477P.
PR	14-AUG-2000; 2000US-0225758P.			PR	08-NOV-2000;	2000US-0246478P.
PR	14-AUG-2000; 2000US-0225759P.			PR	08-NOV-2000;	2000US-0246523P.
PR	14-AUG-2000; 2000US-0225758P.			PR	08-NOV-2000;	2000US-0246524P.
PR	14-AUG-2000; 2000US-0225759P.			PR	08-NOV-2000;	2000US-0246525P.
PR	14-AUG-2000; 2000US-0225758P.			PR	08-NOV-2000;	2000US-0246526P.
PR	14-AUG-2000; 2000US-0225759P.			PR	08-NOV-2000;	2000US-0246527P.
PR	14-AUG-2000; 2000US-0225758P.			PR	08-NOV-2000;	2000US-0246528P.
PR	14-AUG-2000; 2000US-0225759P.			PR	08-NOV-2000;	2000US-0246532P.
PR	14-AUG-2000; 2000US-0225758P.			PR	08-NOV-2000;	2000US-0246609P.
PR	14-AUG-2000; 2000US-0225759P.			PR	08-NOV-2000;	2000US-0246610P.

QY 973 CATGAAACAGGACCAAGATAAACAATCTCATCTTTGACGAGTTTGGACATATACACCCG 1032
Dd |||||||
QY 323 HisGluThrGlyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsnPro 342
Dd |||||||
QY 1033 GAAAGAACCATCACTGTGAAGGGCACAGTTGAGGCTGTGCGAGTGTGAGATAGAGATT 1092
Dd |||||||
QY 343 GluArgThrIleThrValLysGlyThrValGluAlaCysAlaSerAlaGluIleGluIle 362
QY 1093 ATGAGAAGCTGCTGAGGCGCTTTGAAATGATATGCTGGCTGTAAACCAACAGCCAAAT 1152
Dd |||||||
QY 363 MetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaValAsnGlnGlnAlaAsn 382
QY 1153 CTGATCCAGGCTTGAACCTCAGCGCACTTGGCATCTTTCAACAGGACTGTCCTGCTA 1212
Dd |||||||
QY 383 LeuIleProGlyLeuAsnLeuSerAlaLeuGlyIlePheSerThrGlyLeuSerValLeu 402
QY 1213 TCTCCACAGCAGGCGCCCGGAGCTCCCGCCCTGCCCTACACACCCCTTCCACTACC 1272
Dd |||||||
QY 403 SerProProAlaGlyProArgGlyAlaProProAlaAlaProTyrHisProPheThrThr 422
QY 1273 CACTCCGGATCTCTCCAGCTGTACCCCATCACAGTTTGGCCCGCTTCCCGCATCAT 1332
Dd |||||||
QY 423 HisSerGlyTyrPheSerSerLeuTyrProHisGlnPheGlyProPheProHisHis 442
QY 1333 CACTCTTATCCAGACAGGAGATTGTGAATCTCTTCATCCCAACCCAGGCTGTGGCGCC 1392
Dd |||||||
QY 443 HisSerTyrProGluGlnGluIleValAsnLeuPheIleProThrGlnAlaValGlyAla 462
QY 1393 ATCATCGGGAAGAGGGGGCACACATCAACAGCTGGCGAGATTCCGCGGAGCTCTATC 1452
Dd |||||||
QY 463 IleIleGlyLysGlyAlaHisIleLysGlnLeuAlaArgPheAlaGlyAlaSerIle 482
QY 1453 AGATTCCCTCGGAGGCCAGAGCTGAGGAAAGGATGTCATCATCACCGGCCCA 1512
Dd |||||||
QY 483 LysIleAlaProAlaGluGlyProAspValSerGluArgMetValIleIleThrGlyPro 502
QY 1513 CCGAAGCCAGTTTCAAGGCCAGGAGCGATCTTTGGGAACCTGAAAGAGGAAACTTC 1572
Dd |||||||
QY 503 ProGluAlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluGluAsnPhe 522
QY 1573 TTTAACCCCAAGAAAGTGAAGCTGGAAGCGCATATCAGAGTGCCTCTTCCACAGCT 1632
Dd |||||||
QY 523 PheAsnProLysGluGluValLysLeuGlu**HisIleArgValProSerSerThrAla 542
QY 1633 GCGCGGTGATTGGCAAGTGGCAAGACCGTGAACCACTGCAGAACTTAAACAGTGA 1692
Dd |||||||
QY 543 GlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsnLeuThrSerAla 562
QY 1693 GAAGTCACTGCTGCTGACCAAAACCGCAGATGAAATCAGGAAGTGCATCGTCAGATT 1752
Dd |||||||
QY 563 GluValIleValProArgAspGlnThrProAspGluAsnGluGluValIleValArgIle 582
QY 1753 ATCGGCACTCTTTGTACCAAGTGCAGCGCAAGATCAGGGAATTTGACACAG 1812
Dd |||||||
QY 583 IleGlyHisPheAlaSerGlnThrAlaGlnArgLysIleArgGluIleValGlnGln 602
QY 1813 GTNAGCAGCAGGAGCAGAATACCTTCAGGAGTGCCTCAGCCGACAG 1866
Dd |||||||
QY 603 ValLysGlnGlnGluGlnLysTyrProGlnGlyValAlaSerGlnArgSerLys 620
RESULT 3
ABU55232
ID ABU55232 standard; protein; 620 AA.
XX AC ABU55232;
XX DT 18-MAR-2003 (first entry)
XX DE Human novel polypeptide #319.
XX KW Human; neural disorder; immune system disorder; renal disorder;
KW muscular disorder; respiratory disease; reproductive disorder;

KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
KW hyperproliferative disorder; inflammatory disease; allergic reaction;
KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
KW haemostatic; antiarteriosclerotic.
OS Homo sapiens.
XX US2002132753-A1.
XX PD 19-SEP-2002.
XX PF 17-JAN-2001; 2001US-00764864.
XX PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214886P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 23-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX Rosen CA, Ruben SM, Barash SC;
PI WPI; 2003-147444/14.
XX N-PSDB; ABX73491.

XX New polypeptides and nucleic acids, useful in gene therapy for treating,
PT inhibiting or preventing e.g. neural, immune system, muscular,
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
PT renal disorders.
XX Claim 11; SEQ ID NO 1116; 402pp; English.

XX The invention relates to human novel polypeptides and their associated
CC polynucleotides. The polypeptides and polynucleotides are useful in gene
CC therapy for treating, inhibiting or preventing neural disorders, immune
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
CC infarction) and cancerous diseases. Sequences ABUS4914-ABUS5699 and
CC ABUS5748 represent human novel polypeptides of the invention

SQ Sequence 620 AA;

Alignment Scores:
Pred. No.: 2,146-268 Length: 620
Score: 3065.50 Matches: 603
Percent Similarity: 97.73% Conservative: 1
Best Local Similarity: 97.57% Mismatches: 11
Query Match: 50.08% Indels: 3
DB: 6 Gaps: 1

US-09-270-437D-6 (1-3412) x ABUS5232 (1-620)

QY 22 CGCCGGTACCGCGCGCGGAGCGCGGCTCTCGGGAGAGACGCG-----ATG 72
DB 3 ArgArgTyrAlaCysArgTyrArgSerGlyIleProGlySerThrHisAlaSerGlyMet 22
QY 73 ATGAACAAGCTTATCATCGGNAACCTGAGCCCGCGCGTACCGCGGAGACCTCCGGGAG 132
DB 23 MetAsnLysLeuTyrIleGlyAsnLeuSerProAlaValThrAlaAspLeuArgGln 42
QY 133 CTCCTTTGGGAGCAGGAAGTCCCTCTGCGGAGACAGTCTGCTCAAGTCCGGCTACGCC 192
DB 43 LeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLysSerGlyTyrAla 62
QY 193 TTCGTGGACTACCCGACAGAACTGGGCCATCGCGCCATCGAGACCTCTCGGGTAAA 252
DB 63 PheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGluThrLeuSerGlyLys 82
QY 253 GTGGAAATGCATGGGAAATCATGGAAGTTGATGACTCAGTCTCTAAAAGCTAAGGAGC 312
DB 83 ValGluLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLysLeuArgSer 102
QY 313 AGGAAATTCAGATTTCGAAACATCCCTCTCACCTGCGAGTGGAGCTGTGGATGCACTT 372
DB 103 ArgLysIleGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspGlyLeu 122
QY 373 TTGGTCAATATGGACAGCTGGAGAAATGTGGAACAAGTCAACACAGACACAGAAACCGCC 432
DB 123 LeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThrAspThrGluThrAla 142
QY 433 GTTGTCAAGCTCACATATGCACACAGACAGAGACAAATAGCCATCGGAGCTAAGC 492
DB 143 ValValAsnValThrTyrAlaThrArgGluGluAlaLysIleAlaMetGluLysLeuSer 162
QY 493 GGGCATCAGTTTGAACAATCTCTTCAAGATTCTCTACATCCCGGATGAAGAGTGAGC 552
DB 163 GlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIleProAspGluGluValSer 182
QY 553 TCCCTCTGCCCGCTCAGGAGCGCCAGCGGTGGGGACCACTCTTCCCGGAGCAAGGCCAC 612

DB 183 SerProSerProGlnArgAlaGlnArgGlyAspHisSerSerArgGluGlnGlyHis 202
QY 613 GCCCTTGGGGGCACTTCTCAGGCCAGACAGATTGATTTCCCGCTGCGGATCTTGTCCTCC 672
DB 203 AlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeuValPro 222
QY 673 ACCCAGTTTGTGTGGCCATCATCGAAAGAGGAGGCTTGACCAATAAAGAACATCACTAAG 732
DB 223 ThrGlnPheValGlyAlaIleGlyLysGluGlyLeuThrIleLysAsnIleThrLys 242
QY 733 CAGACCCAGTCCCGGTAGATATCATAGAAAAGAGAAGTCTGGAGCTGCAGAGAGCCT 792
DB 243 GlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGlyAlaAlaGluLysPro 262
QY 793 GTCACCATCATGCCACCCAGAGGGGACTTCTAGAGATCCCGATCCCGATGATTTCTGAATC 852
DB 263 ValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeuGluIle 282
QY 853 ATGCAGAAAGAGGAGGATGACCAAACTAGCCGAAGAGATTCCTCTGAAAATCTTGCA 912
DB 283 MetGlnLysGluAlaAspGluThrLysLeuAlaGluGluIleProLeuLysIleLeuAla 302
QY 913 CACAATGGCTTGTGTGAAGACTGATTGGAAGAGGAGGAGCAAAATTTGAAGAAATTTGAA 972
DB 303 HisAsnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLysLysIleGlu 322
QY 973 CATGAACAGGAGGACCAAGATTAACATCTCATCTTTGAGGATTTGAGATATACACCCG 1032
DB 323 HisGluThrGlyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsnPro 342
QY 1033 GAAAGAACCATCACTGTGAAGGCGACAGTTGAGGCTGTGCGAGTGTGAGATAGAGATT 1092
DB 343 GluArgThrIleThrValLysGlyThrValGluAlaCysAlaSerAlaGluIleGluIle 362
QY 1093 ATGAAGAAGCTGCTGAGGCTTTGAAATGATGCTGGTGTGTTAAACCAACAGCCCAAT 1152
DB 363 MetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaValAsnGlnAlaAsn 382
QY 1153 CTGATCCAGAGGTGAACTCAGCGCACTTGGCATCTTTCAACAGGAGTGTCCGTGCTA 1212
DB 383 LeuIleProGlyLeuAsnLeuSerAlaLeuGlyIlePheSerThrGlyLeuSerValLeu 402
QY 1213 TCTCACAGCAGGCGCCCGCGAGCTCCCGCGCTGCCCTTACACCCCTTCACTACC 1272
DB 403 SerProProAlaGlyProArgGlyAlaProProAlaAlaProTyrHisProPheThrThr 422
QY 1273 CACTCCGGATCTTCTCAGCCTGTACCCCATCACAGTTTGGCCCGTTCCTCCGATCAT 1332
DB 423 HisSerGlyTyrPheSerSerLeuTyrProHisGlnPheGlyProPheProHisHis 442
QY 1333 CACTCTTATCCAGAGCAGGAGATTGTGAATCTTTCATCCCAACCCAGCGTGTGGCGCC 1392
DB 443 HisSerTyrProGluGlnGluIleValAsnLeuPheIleProThrGlnAlaValGlyAla 462
QY 1393 ATCATCGGAAAGAGGGGCGCACACATCAACAGCTGGCGAGATTTCGCGGAGCCCTCTATC 1452
DB 463 IleIleGlyLysLysGlyAlaHisIleLysGlnLeuAlaArgPheAlaGlyAlaSerIle 482
QY 1453 AAGATTGCCCTCGGAAAGGCCAGAGCTCAGCGAAGAGGATGGTTCATCATCCCGGCGCA 1512
DB 483 LysIleAlaProAlaGluGlyProAspValSerGluArgMetValIleIleThrGlyPro 502
QY 1513 CCGGAAGCCCATGTTCAAGGCCAGGAGCGGATCTTTGGGAACTTGAAGAGAGAAATTC 1572
DB 503 ProGluAlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluGluAsnPhe 522
QY 1573 TTTAAACCCCAAGAGAGTGAAGCTGGAAGCGCATATCAGAGTGCCTCTTTCACAGCT 1632
DB 523 PheAsnProLysGluGluValLysLeuGlu***HisIleArgValProSerSerThrAla 542
QY 1633 GGCGGGTGTGTCGAAAGGTGGCAAGCCCTGAACGAACTGCAGAACTTAACAGTGA 1692

Db 543 GlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsnLeuThrSerAla 562
QY 1693 GAAGTCATCTGCTCTGTGACCAAAACCCAGATGAAATCAGGAATGATCGTCAGAATT 1752
Db 563 GluValIleValProArgAspGlnThrProAspGluAsnGluGluValIleValArgIle 582
QY 1753 ATCGGGCACTCTTGTCTAGCCAGACTGCACAGCGGAGATCAGGGAAATGTACACAG 1812
Db 583 IleGlyHisPheIleAlaSerGlnThrAlaGlnArgLysIleArgGluIleValGlnGln 602
QY 1813 GTGAAGCAGCAGGACGAGAATACCTCTCAGGGAGTCGCCCTCACAGCGCAGCAAG 1866
Db 603 ValIysGlnGlnGluGlnLysTyrProGlnGlyValAlaSerGlnArgSerLys 620

RESULT 4
ABG06795
ID ABG06795 standard; protein; 594 AA.

XX ABG06795;
XX
XX
DT 13-FEB-2002 (first entry)
DE Novel human diagnostic protein #6786.
DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
KW Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
PI
PI WPI; 2001-639362/73.
DR N-PSDB; AAS70982.
XX

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX

XX Claim 20; SEQ ID NO 37154; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 594 AA;
Alignment Scores:
Pred. No.: 5,51e-255 Length: 594
Score: 2917.50 Matches: 577
Percent Similarity: 93.08% Conservative: 1
Best Local Similarity: 92.91% Mismatches: 0
Query Match: 47.66% Indels: 43
DB: 4 Gaps: 1

US-09-270-437D-6 (1-3412) x ABG06795 (1-594)

QY 4 AGCGAGGAGGCGCAGGAGCCCGGTACCGGGCGGGGAGCCGCGGTCTCTCGGGAAG 63
Db 17 AsnGlyGlyGlyGluGluArgArgValProGlyArgGlySerArgGlyLeuSerGlyLys 36
QY 64 AGACGGATGATGAACAAGCTTTACATCGGGAACCTGAGCCCGCCCTCACCGCGGAC 123
Db 37 ArgArgMetMetAsnLysLeuTyrIleGlyAsnLeuSerProAlaValThrAlaAsp 56
QY 124 CTCGGCAGCTCTTTGGGACAGGAAGCTGCCCTCGCGGACAGGTCTCTGCTGAAGTCC 183
Db 57 LeuArgGlnLeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLeuLysSer 76
QY 184 GGCTAGCCTTCTGTGACTACCCCGACACAGAACTGGGCCCATCCGCGCATCGAGACCTC 243
Db 77 GlyTyrAlaPheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGluThrLeu 96
QY 244 TCGGGTAAAGTGAATTGCATGGGAAATCATCGAAGTTGATTACTCAGTCTCTAAAAAG 303
Db 97 SerGlyLysValGlnLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLys 116
QY 304 CTAAGGAGCAGGAAAAATTCAGATTTCGAAACATCCCTCCTCACCTGAGTGGAGGTGTG 363
Db 117 LeuArgSerArgLysIleGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeu 136
QY 364 GATGGACTTTTGGCTCAATATGGGACAGTGGAGAAATGTGGACAAGTCAACACAGACACA 423
Db 137 AspGlyLeuLeuAlaGlnTyrGlyThrValGluAsnValGlnValAsnThrAspThr 156
QY 424 GAAACCGCCGTTCTCAACGTTCACATATGCAACAGAGAGAGAAACAAAATAGCATCGAG 483
Db 157 GluThrAlaValValAsnValThrTyrAlaThrArgGlnGluAlaLysIleAlaMetGlu 176
QY 484 AAGCTAAGCGGGCATCAGTTTGAGAACTACTCTTCAAGATTTCCTACATCCCGGATGAA 543
Db 177 LysLeuSerGlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIleProAspGlu 196
QY 544 GAGTGAGTCTCCCTTCGCCCCCTCAGCGAGCCCGAGCGTGGGACCACTCTTCCCGGGAG 603
Db 197 GluValSerSerProSerProGlnArgAlaGlnArgGlyAspHisSerSerArgGlu 216
QY 604 CAAGGCCACGCCCTCGGGGSCACTTCTCAGGCCCAGACAGATTGATTTCGCGCTCGGATC 663
Db 217 GlnGlyHisAlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeuArgIle 236
QY 664 CTGGTCCCCACCCAGTTTGTGTGTCATCATCGAAAGGAGGGCTTGACCATAAAGAAC 723
Db 237 LeuValProThrGlnPheValGlyAlaIleGlyLysGluGlyLeuThrIleLysAsn 256
QY 724 ATCACTAAGCAGACCCAGTCCCGGTAGATATTCATAGAAAAGAGAACTCTGTGAGCTGCA 783
Db 257 IleThrLysGlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGlyAlaAla 276
QY 784 GAGAAGCTGTCAACCATCCATCCATCCACCCAGAGGGGACTTCTGAAGCATGCCCATGATT 843
Db 277 GluLysProValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgMetIle 296
QY 844 CTTTAAATCATGCAGAAAGAGGAGATGAGACCAAACTAGCCGGAAGAGATTCTCTCTGAA 903
Db 297 LeuGluIleMetGlnLysGluAlaAspGluThrLysLeuAlaGluGluIleProLeuLys 316

Alignment Scores:

Pred. No.: 6,19e-245 Length: 556
Score: 2806.50 Matches: 556
Percent Similarity: 92.82% Conservative: 0
Best Local Similarity: 92.82% Mismatches: 0
Query Match: 45.85% Indels: 43
DB: 5 Gaps: 1

US-09-270-437D-6 (1-3412) x ABG96346 (1-556)

```
QY 70 ATGATGAACAGCTTTACATCGGGAACCTGAGCCCGCGCTGACCGCGAGCCTCCGG 129
DB 1 MetMetAsnLysLeuYrilleGlyAsnLeuSerProAlaValThrAlaAspLeuArg 20
QY 130 CAGCTCTTTGGGACAGGAAGCTGCCCTCGCGGGCAGAGTCTCTGCTGAAGTCGGGTAC 189
DB 21 GlnLeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLysSerGlyTyr 40
QY 190 GCCTTCTGTGACTACCCCGACAGAACTGGGCCATCCGCGCCATCGAGACCCTCTCGGGT 249
DB 41 AlaPheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGluThrLeuSerGly 60
QY 250 AAGTGTGAATTGCATGTGGAAATCATGCAAGTTGATTACTCAGTCTCTAAAAAGCTAAGG 309
DB 61 LysValGlnLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLysLeuArg 80
QY 310 AGCAGGAAATTCAGATTCGAAACATCCCTCCTCACCTGAGTGGGAGGTGTTGGATGGA 369
DB 81 SerArgLysIleGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspGly 100
QY 370 CTTTTGGCTCAATATGGGACAGTGGAGAATGTGGAACAAGTCAACACAGACAGAAACC 429
DB 101 LeuLeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThrAspThrGluThr 120
QY 430 GCCCTGTCTCAAGTCACATATGCAACAGAGAAGCAAAAAATAGCCATGGAGAAGCTA 489
DB 121 AlaValValAsnValThrTyrAlaThrArgGluGluAlaLysIleAlaMetGluLysLeu 140
QY 490 AGCGGGCATCAGTTTGAGAACTACTCCTTCAAGATTTCTTACATCCCGGATGAGAGGTG 549
DB 141 SerGlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIleProAspGluGluVal 160
QY 550 AGCTCCCTTCGCCCTCAGCGAGCCAGCGTGGGACCACTCTTCCCGGGAGCAGGC 609
DB 161 SerSerProSerProProGlnArgAlaGlnArgGlyAspHisSerSerArgGluGlnGly 180
QY 610 CACGCCCTTGGGGCACTTCTCAGGCCAGACAGATTGATTTCCCGCTGCGGATCCTGGTC 669
DB 181 HisAlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeuVal 200
QY 670 CCACCCAGTTTGTGTGGCATCATCGGAAGAGGGCTTGACCATAAAGAACATCACT 729
DB 201 ProThrGlnPheValGlyAlaIleIleGlyLysGluGlyLeuThrIleLysAsnIleThr 220
QY 730 AAGCAGACCCAGTCCCGGTAGATATCCATAGAAAGAGAACTCTGAGCTGCAGAGAAG 789
DB 221 LysGlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGlyAlaAlaGluLys 240
QY 790 CTTGTCTACCATCATGATGCCACCCAGAGGGGACTTCTGAAGCATGCCGATGATTCTTGAA 849
DB 241 ProValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeuGlu 260
QY 850 ATCATGCAAGAAGGAGGATGAGACCAAACTAGCCGAAGAGATTCTCTGAAATCTTG 909
DB 261 IleMetGlnLysGluAlaAspGluThrLysLeuAlaGluGluIleProLeuLysIleLeu 280
QY 910 GCACACAAATGCTTGGTGGAGACTGATTGGAAAAGAGCAGCAAAATTTGAAGAAATT 969
DB 281 AlaHisAsnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLysLysIle 300
QY 970 GAACATGAACAGGGACCAAGATAACAATCTCATCTTTTGAGGATTTGAGCATATCAAC 1029
```

```
DB 301 GluHisGluThrGlyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsn 320
QY 1030 CCGAAAGAACCATCATCTGTGAAGGCGACAGTGTGAGCCCTGTGCCAGTGTGTGAGATAGAG 1089
DB 321 ProGluArgThrIleThrValLysGlyThrValGluAlaCysAlaSerAlaGluIleGlu 340
QY 1090 ATTATGAAGAAGCTGCTGAGGCTTTGAAAAATGATATGTGCTGTGTACCAACAAGCC 1149
DB 341 IleMetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaValAsn----- 357
QY 1150 AATCTGATCCACAGGTTGAACCTCAGCGCACTTGGCATCTTTTCAACAGGACTGTCCGTG 1209
DB 357 ----- 357
QY 1210 CTATCTCCACAGAGGCGCCCGGAGGTCCCCCGCTGCCCCCTACCAACCCCTTCACT 1269
DB 357 ----- 357
QY 1270 ACCACTCCGGATACTTCTCCAGCCTGTACCCCATCACCAGTTTGGCCCGTTCGCCGAT 1329
DB 358 ThrHisSerGlyTyrPheSerSerLeuTyrProHisGlnPheGlyProPheProHis 377
QY 1330 CATCACTCTTATCCAGAGCAGGAGATTGTGAATCTCTTCATCCCAACCCAGGCTGTGGC 1389
DB 378 HisHisSerTyrProGluGlnGluIleValAsnLeuPheIleProThrGlnAlaValGly 397
QY 1390 GCATCATCGGGAAGGGGGCACACATCAACAGCTGCGGAGATTGCCGAGCCTCT 1449
DB 398 AlaIleIleGlyLysLysGlyAlaHisIleLysGlnLeuAlaArgPheAlaGlyAlaSer 417
QY 1450 ATCAAGATTCCCTCGGAGGCGCCAGAGTCCAGCGTCCAGGAGGATGTCATCATCAGCGG 1509
DB 418 IleLysIleAlaProAlaGluGlyProAspValSerGluArgMetValIleIleThrGly 437
QY 1510 CCACCGGAAGCCAGATTCAAGGCCAGGGACGATCTTTGGGAAACTGAAAGAGGAAAC 1569
DB 438 ProProGluAlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluGluAsn 457
QY 1570 TTTCTTTAAACCCAAAGAAAGTGAAGCTGGAAGCGCATATCAGAGTGCCTCTTCCACA 1629
DB 458 PhePheAsnProLysGluGluValLysLeuGluAlaHisIleArgValProSerSerThr 477
QY 1630 GCTGGCGGGTGAATTGGCAAGGTGGCAGACCGTGAACCAACTGACAGACTTAACCACT 1689
DB 478 AlaGlyArgValIleGlyLysGlyLysGlyLysThrValAsnGluLeuGlnLeuThrSer 497
QY 1690 GCAGAACTCATCTGCTGCTCGTACCAAAACCCAGATGAAATGAGGAAGTGCATCGTCA 1749
DB 498 AlaGluValIleValProArgAspGlnThrProAspGluAsnGluGluValIleValArg 517
QY 1750 ATTATCGGCACTTCTTTTGTGCTGACAGACTGCACAGCGCAAGATCAGGAAATTTGTA 1809
DB 518 IleIleGlyHisPhePheAlaSerGlnThrAlaGlnArgLysIleArgGluIleValGln 537
QY 1810 CAGGTGACGACGAGGACAGCAATACCTCCAGGGAGTCCCTCAGCCAGCAGCAAG 1866
DB 538 GlnValLysGlnGlnGluGlnLysTyrProGlnGlyValAlaSerGlnArgSerLys 556
RESULT 6
ABU89799
ID ABU89799 standard; protein; 555 AA.
XX ABU89799;
AC
XX 10-JUL-2003 (first entry)
DT
XX
XX Novel human protein NOV14a.
DE
XX Human; cytostatic; DAPK3-Agonist; DAPK3-Antagonist; cancer; NOV.
XX Homo sapiens.
OS
XX
XX W02003031571-A2.
PN
```

XX PD 17-APR-2003.
 XX PF 02-OCT-2002; 2002WO-US031357.
 XX PR 05-OCT-2001; 2001US-0327454P.
 PR 09-OCT-2001; 2001US-0327917P.
 PR 09-OCT-2001; 2001US-0328029P.
 PR 09-OCT-2001; 2001US-0328056P.
 PR 12-OCT-2001; 2001US-0328849P.
 PR 15-OCT-2001; 2001US-0329414P.
 PR 17-OCT-2001; 2001US-0330142P.
 PR 22-OCT-2001; 2001US-0341058P.
 PR 24-OCT-2001; 2001US-0343629P.
 PR 29-OCT-2001; 2001US-0349575P.
 PR 01-NOV-2001; 2001US-0346357P.
 PR 25-JUN-2002; 2002US-0391342P.
 PR 01-OCT-2002; 2002US-0026244S.
 XX PR (CURA-) CURAGEN CORP.
 XX PI Alsbrook JP, Burgess CE, Catterton E, Chant JS, Chaudhuri A;
 PI Edinger SR, Gerlach VL, Giot L, Gorman L, Guo X, Kekuda R;
 PI Mezes PS, Millet I, Ooi CE, Patturajan M, Rieger DK, Spytek KA;
 PI Taupier RJ, Zernhusen BD, Zhong H, Zhong M;
 XX DR WPI; 2003-381704/36.
 DR N-PSDB; ACA90176.
 XX PR New DAPK3 polypeptide, useful for preparing a composition for treating or
 PT preventing e.g., cancer.
 XX PS Claim 2; Page 129; 253pp; English.
 CC The invention describes an isolated polypeptide comprising any of 33 90-
 CC 1273 amino acid sequences (I) given in the specification or its mature
 CC form, a sequence that is at least 95 % identical to (I), or a sequence
 CC comprising one or more conservative substitutions in the amino acid
 CC sequence of (I). The polypeptide is useful for preparing a composition
 CC for treating or preventing e.g. cancer. This is the amino acid sequence
 CC of a novel human NOV protein
 XX SQ Sequence 555 AA;
 Alignment Scores:
 Pred. No.: 6, 01e-222 Length: 555
 Score: 2553.00 Matches: 507
 Percent Similarity: 87.81% Conservative: 19
 Best Local Similarity: 84.64% Mismatches: 29
 Query Match: 41.71% Indels: 44
 DB: 6 Gaps: 2
 US-09-270-437D-6 (1-3412) x ABU89799 (1-555)
 QY 70 ATGATGAACAGCTTTACATCGGAACTGTAGCCCGCCGCTCACCGCGACGACCTCGG 129
 Db 1 MetMetAsnLysLeuPheIleGlyAsnLeuSerProAlaValThrAlaGluAspLeuArg 20
 QY 130 CAGCTCTTTGGGACAGGAAGCTGCCCTGGCGGACAGGTCCTGCTGAAGTCCGGCTAC 189
 Db 21 GlnLeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLysSerArgTyr 40
 QY 190 GCCTTCGTGGTACTCCCGACAGAACTGGGCCATCCCGGCATCGAGACCTCTCGGGT 249
 Db 41 AlaPheValAspTyrProAspGlnAsnTrpAlaIleArgThrIleGluThrLeuSerGly 60
 QY 250 AAATGGGAATGCATGGGAAATCATGGAATTCATTACTCAGTCTCTAAAAGCTAAGG 309
 Db 61 GlnValGluLeuHisGlyLysIleMetGluValAspTyrSerValSerIleLysLeuArg 80
 QY 310 AGCAGAAATTCAGATTCGAAACATCCCTCCTCACCTGCAGTGGGAGGTGTGGATGGA 369
 Db 81 SerArgAsnIleProIleArgAsnIleProHisLeuGlnTrpGluValLeuAspGly 100

QY 370 CTTTGGCTCAATATGGGACAGTGTGGAACAGTCTGAACAGTCAACACACACACACACAC 429
 Db 101 LeuLeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThrAspThrGluThr 120
 QY 430 GCCGTTGTCAACAGTCAACATATGCAACAGAGAAAGCAAAATAGCCATGAGAGCTA 489
 Db 121 AlaValValAsnValThrTyrAlaThrLysGluValLysIleAlaMetLysLysLeu 140
 QY 490 AGCGGCGATCAGTTTGAGAACTATCTCTTCAAGATTCTCTACATCCCGATGAAGGTG 549
 Db 141 SerGlyHisGlnPheGluAsnHisTyrPheLysIleSerTyrIleProAspAspGluVal 160
 QY 550 AGCTCCCTCCGCGCCCTCAGCGCCAGCGTGGGAGCCCTCTTCCCGGAGCAAGGC 609
 Db 161 SerCysProSerProGlnArgAlaGlnArgGlyAspHisSerSerTyrGluGlnGly 180
 QY 610 CACGCCCTGGGGCAGCTTCTCAGGCCAGACAGATTGATTTCCCGCTCGGATCCTGTC 669
 Db 181 GlnAlaProGlyGlySerSerGlnAlaArgGlnIleAspPheProLeuArgValLeuPhe 200
 QY 670 CCCACCCAGTTGTGTGTGTCATCGAAAGAGGCGCTTGACCATAAAGACATCACT 729
 Db 201 ProThrGlnPheValGlyAlaIleGlyLysGluGlyLeuThrIleLysAsnIleThr 220
 QY 730 AAGCAGACCCAGTCCCGGTAGATATCCATAGAAAGAACTCTGGAGCTGCAGAGAG 789
 Db 221 LysGlnSerArgSerArgValAspIleTyrArgGlnGlnAsnSerArgAlaGluLys 240
 QY 790 CTTGTCACTCCATCCACCCAGAGGGGAGCTTCTGAAGCATGCCGATGATCTTGA 849
 Db 241 ProValThrMetHisAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeuGlu 260
 QY 850 ATCATGCAAGAGGCGAGATGACCAACTAGCCGAAAGAGATTCCTCTGAAATCTTG 909
 Db 261 IleMetGlnLysGluAlaAspGluAlaLysLeuAlaGluGluIleProLeuLysIleLeu 280
 QY 910 GCACCAATGCTTGTGTGGAAGACTGATTCGAAAGAAAGCAAGAAATTTCAAGAAAT 969
 Db 281 AlaHisGlnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLysAsn 300
 QY 970 GAACATGAACACAGGACCAAGATAACATCTCTTTCAGGAGTTTGAGCATATACAA 1029
 Db 301 GluHisGluThrGlyThrLysIleThrIleSerSerSerGlnAspLeuSerIleTyrAsn 320
 QY 1030 CCGAAAGAACCATCACTGTCAAGGACAGTTCAGGCGCTGTGCGAGTGTGAGATAGAG 1089
 Db 321 ProGluArgThrIleThrValLysGlyThrValGluValCysAlaSerAlaGluIleGlu 340
 QY 1090 ATTATGAAGAGCTCGCTGAGGCGCTTTGAAATATATATGCTGGCTGTGTTAACCA 1149
 Db 341 IleMetLysLysLeuArgGluAlaPheGluAsnAspThrLeuThrValAsn 357
 QY 1150 AATCTGATCCAGGGTTGAACCTCAGCGCACTGGCATCTCTTTCAACAGGAGTGTCCGTG 1209
 Db 357 ----- 357
 QY 1210 CTATCTCCACAGAGGCGCCCGGAGAGTCCCCCGCTGCCCTTACCACCCCTTCACT 1269
 Db 357 ----- 357
 QY 1270 ACCCACTCCGATATCTTCTCCAGCGCTGTACCCCATCACAGTTTGGCCGCTTCCCGCAT 1329
 Db 358 ThrHisPheGlyTyrPheSerSerLeuTyrProHisArgGlnPheGlyProPheProHis 377
 QY 1330 CATCACTCTTATCCAGAGCAGGAGATGTGATCTCTTCATCCCAACCCAGGCTGTGGC 1389
 Db 378 HisHisSerTyrProGluGlnGluIleValAsnLeuPheIleProThrGlnGlyValGly 397
 QY 1390 GCCATCATCGGAAAGAGGGGGCACATCAACAGCTGGGAGATTCGCGGAGCTCT 1449
 Db 398 AlaIleIleGlyLysLysGlyAlaHisIleLysGlnLeuAlaArgPheValGlyAlaSer 417

QY 1450 ATCAAGATTGCCCTCGGAGGCCACAGCTCAGCGAAGGATGGTTCATCATCACCGGG 1509
 Db ||||||| ||||||| ||| : : : ||| |||||||
 418 IlelyslleAlaProAlaArgSerPro--LeuArgGlnArgLysValIleIleThrTrp 436
 QY 1510 CCACCGAAGCCAGTTCAAAGGCCAGGCGGATCTTTGGGAACTGAAAGAGGAAAC 1569
 Db ||||||| : : : ||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 437 ProProGluSerGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluGluAsn 456
 QY 1570 TTCTTTAACCCCAAGAGAGTGAAGTGAAGCGCATATCAGAGTGCCCTCTCCACA 1629
 Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 457 PhePheAsnProLysGluAspValLysLeuGluThrHisIleArgValProSerSerThr 476
 QY 1630 GCTGGCGGGTGTGTCGCAAGGCGCAAGACCTGAACGACTGACGAACTTGAACAGT 1689
 Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 477 AlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGluAsnLeuIleSer 496
 QY 1690 GCAGAAGTCATCGTCCTGTGACCAACGCCAGATGAAATAGAGGAAGTGATCGTCAGA 1749
 Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 497 AlaGluValIleValProArgAspGlnThrProAspGluAsnGluMetIleValArg 516
 QY 1750 ATTATCGGGGACTTCTTTGCTAGCCAGACTGCACAGCGCAAGATCAGGGAAATGTACAA 1809
 Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 517 IleIleGlyHisPheAlaSerGlnThrAlaGlnArgLysIleArgGluIleValGln 536
 QY 1810 CAGGTGAAGCAGCAGGAGCAGAAATACCTCAGGAGTTCGCTCACAGCGCAGCAAG 1866
 Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 537 GlnValLysGlnGlnGlnLysTyProGlnGlyValAlaSerGlnArgSerLys 555

RESULT 7

AAY30649
 ID AAY30649 standard; protein; 577 AA.
 AC AAY30649;
 XX DT 17-NOV-1999 (first entry)
 XX A murine c-myc coding region determinant binding protein.
 DE
 XX
 XX c-myc coding region determinant binding protein; CRD-BP; tumor; c-myc;
 KW endonucleolytic attack; half-life; breast cancer; colon cancer;
 KW pancreatic cancer.
 XX
 XX Mus musculus.
 XX
 XX WO9946594-A2.
 XX
 XX PD 16-SEP-1999.
 XX
 XX 05-MAR-1999; 99WO-US004897.
 XX
 XX 09-MAR-1998; 98US-0077372P.
 XX
 XX (WISC) WISCONSIN ALUMNI RES FOUND.
 XX
 XX Ross J;
 XX
 XX WPI; 1999-551506/46.
 XX N-PSDB; AAZ10617.
 XX

Diagnosing presence or absence of a tumor in a human by examining c-myc
 coding region determinant-binding protein.

Example; Fig 1A-D; 79pp; English.

The present sequence represents a murine c-myc coding region determinant
 binding protein (CRD-BP). The presence or absence of a tumor can be
 determined by determining the levels of CRD-BP present in the suspect
 tissue, where the CRD-BP shields c-myc RNA from endonucleolytic attack
 and so prolongs its half-life. The methods are used for diagnosing
 presence or absence of a tumor in a human, especially breast, colon and
 pancreatic cancer. They are also used to inhibit cancer cell growth

Sequence 577 AA;

Alignment Scores:
 Pred. No.: 1.47e-173 Length: 577
 Score: 2019.50 Matches: 394
 Percent Similarity: 78.94% Conservative: 82
 Best Local Similarity: 65.34% Mismatches: 96
 Query Match: 32.99% Indels: 31
 DB: 2 Gaps: 4

US-09-270-437D-6 (1-3412) x AAY30649 (1-577)

QY 73 ATGACAAAGCTTTACATCGGGAACCTCAGCCGCCCTCAGCCGACGACCTCCGGCAG 132
 Db 1 MecAsnLysLeuTyrlleGlyAsnLeuAsnGluSerValThrProAlaLeuGluLys 20
 QY 133 CTCTTTGGGACAGGAAGCTGCCCTCGGCGGACAGCTCCTGCTGAAGTCCGGGTACGCC 192
 Db 21 ValPheAlaGluHisIleSerTyrlleSerGlyGlnPheLeuValLysSerGlyTyrlleAla 40
 QY 193 TTCTGTGACTACCCCGACACAGAACTGGGCCATCCGGCCATCGAGACCCCTCTCGGGTAAA 252
 Db 41 PheValAspCysProAspGluHisTrpAlaMetLysAlaIleGluThrPheSerGlyLys 60
 QY 253 GTCGAATTGCATGGGAAATCATGTGAAGTTGATTCTCAGTCTCTAAAACCTTAAGCAGC 312
 Db 61 ValGluLeuGlnGlyLysArgLeuGluMetGluHisSerValProLysLysGlnArgSer 80
 QY 313 AGGAAATTCAGATTTCGAAACATCCCTCTCAGCTCAGTGGGAGGTGTTGGATGGAATT 372
 Db 81 ArgLysIleGlnIleArgAsnIleProProGlnLeuArgTrpGluValLeuAspSerLeu 100
 QY 373 TTGGCTCAATATGGGACAGTGGAGAATGTGGAACAAGTCAACACAGACACAGAAACCGCC 432
 Db 101 LeuAlaGlnTyrlleGlyThrValGluAsnCysGluGlnValAsnThrGluSerGluThrAla 120
 QY 433 GTTGTCAACGTCAATATGCAACAGAGAGAAGCAAAAATAGCCATGGAGAGCTAAGC 492
 Db 121 ValValAsnValThrTyrlleSerAsnArgGluGlnThrArgGlnAlaIleMetLysLeuAsn 140
 QY 493 GGGCATCAGTTTGAAGACTACTCTTCAAGATTTCCTACATCCCGATGAAGAGTGAAGC 552
 Db 141 GlyHisGlnLeuGluAsnHisAlaLeuLysValSerTyrlleProAspGluGlnIleThr 160
 QY 553 TCCCTTTCGCCCTCAGCGAGCCCGAGTGGGACCACTCTTCCCGGGAGAGGCCAC 612
 Db 161 -----GlnGlyProGluAsnGlyArgArgGlyGlyPheGlySerArgGlyGlnProArg 178
 QY 613 -----GCCCTGGGGGCACCTTCTCAGGCCAGACAGATTGATTTCCCGCTG 657
 Db 179 GlnGlySerProValAlaAlaGlyAlaProAlaLysGlnGlnProValAspIleProLeu 198
 QY 658 CGGATCCTGTCCGCCACCCAGTTTGTGTGGTCCCATCATCGGAAGAGGGCTTGACCAT 717
 Db 199 ArgLeuLeuValProThrGlnTyrlleValGlyAlaIleIleGlyLysGluGlyAlaThrIle 218
 QY 718 AAGAACATCACTAAGCAGACCCAGTCCCGGGTGAATATCCATAGAAAAGAGACTCTGGA 777
 Db 219 ArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGly 238
 QY 778 GCTGCGAGAGAGCTGTGCACATCCATGCCACCCAGAGGGGACTTCTGAAGCATGCCG 837
 Db 239 AlaAlaGluLysAlaIleSerValHisSerThrProGlnGlyCysSerSerAlaCysLys 258
 QY 838 ATGATTCTTGAATCATGCAAGAAGAGGAGATGACACCAAACTAGCCGGAAGAGATTCTCT 897
 Db 259 MetIleLeuGluIleMetHisLysGluAlaLysAspThrLysThrAlaAspGluValPro 278
 QY 898 CTGAAATCTTGGCACAACAATGGCTTGGTTGGAAGACTGATTGGAAGAAAGAGCAGAAAT 957
 Db 279 LeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsn 298
 QY 958 TTCAGAAATTCGAACATGAAACAGGACCAACATACATCTCATCTTTGAGGATTG 1017


```
QY 373 TTGGCTCAATATGGACAGTGGAGATGTGGACAAAGTCAACACAGACACAGAAACGGCC 432
Db 101 LeuValGlnThrGlyValValGluSerCysGluGlnValAsnThrAspSerGluThrAla 120
QY 433 GTTGTCAACGTCATATGCAACAAGAGAAAGCAAAATAGCCATGGAGAACTAAAGC 492
Db 121 ValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn 140
QY 493 GGGCATCAGTTTGAAGAACTCTCCTCAAGATTCTTACATCCCGGATGAAGAGTGGAGC 552
Db 141 GlyPheGlnLeuGluAsnPheThrLeuLysValAlaThrIleProAspGluThrAlaAla 160
QY 553 TCCCTTCGCCCTCAG-----CGAGCCACAGCT-----CGGGACCACTCTTCCCGG 600
Db 161 GlnGlnAsnProLeuGlnGlnProArgGlyArgGlyLeuGlnArgGlySerSer 180
QY 601 GAGCAAGGCCACCCCTCGGGGCACCTTCTCAGGCCACAGAGATTGATTTCCCGCTCGCG 660
Db 181 ArgGlnGly---SerProGlySerValSerLysGlnLysProCysAspLeuProLeuArg 199
QY 661 ATCTGTGTCGCCACCCAGTTGTTGTGGCCATCATCGAAAGAGGGCTTGACATAAAG 720
Db 200 LeuLeuValProThrGlnPheValGlyAlaIleGlyLysGluGlyAlaThrIleArg 219
QY 721 AACATCACTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAGAGAACTCTGGAGCT 780
Db 220 AsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGlyAla 239
QY 781 GCAGAGAAGCTGTCAACATCCATGCCACCCAGGGGACTTCTGAAGCATGCCCGCATG 840
Db 240 AlaGluLysSerIleThrLeuSerThrProGluGlyThrSerAlaAlaCysLysSer 259
QY 841 ATTCTTGAATCATGCAAGAAGGCGAGATGACACCAACTAGCCGAGAGATTCTCTCTG 900
Db 260 IleLeuGluIleMethHisLysGluAlaGlnAspIleLysPheThrGluGluIleProLeu 279
QY 901 AAAATCTTGGCACACATGCTTGGTTGAAGACTGATTGAAAAGAGGACAGAAATTTG 960
Db 280 LysIleLeuAlaHisAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsnLeu 299
QY 961 AAGAAATTTGAACATGAACAGGACCAAGATAACAACTCTCATCTTTGCAGAGATTGGAGC 1020
Db 300 LysLysIleGluGlnAspThrAspThrLysIleThrIleSerProLeuGlnGluLeuThr 319
QY 1021 ATATACACCCGGAAGAACATCACTGTGAAGGCGACATTTAGGCGCTGTGCCAGTGCT 1080
Db 320 LeuTyrAsnProGluArgThrIleThrValLysGlyAsnValGluThrCysAlaLysAla 339
QY 1081 GAGATAGAGATTATGAAGAAGTGCGTGAGGCTTTGAAATGATATGCTGGCTGTTAAC 1140
Db 340 GluGluGluIleMethLysLysIleArgGluSerTyrGluAsnAspIleAlaSerMetAsn 359
QY 1141 CAACAAGCCAACTGATCCAGGGTTGAACCTCAGCGCACTTGGCACTTTTCAACAGGA 1200
Db 360 LeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeuGlyLeuPhe----- 376
QY 1201 CTGTCCTGTCTATCTCACACAGAGGGCGCGGAGCTCCCGCTGCCCGCCCTACCAC 1260
Db 377 -----ProProThrSer-----GlyMetProProThr----- 386
QY 1261 CCCTTCACTACCCACTCCGGATCTTCTCCAGCGCTGTACCCCATCACCAGTTTGGCGCG 1320
Db 387 -----SerGlyProProSerAlaMetThrPro-----Pro 396
QY 1321 TTCCCGCATCATCACTCTTATCCAGACGAGGATTTGTGAATCTCTTCAATCCCAACCCAG 1380
Db 397 TyrProGlnPheGluGln---SerGluThrGluThrValHisLeuPheIleProAlaLeu 415
QY 1381 GCTGTGGGCGCATCATCGGAAGAAGGGGCGACACATCAAAAGCTGGCGGAGATTCCGCC 1440
Db 416 SerValGlyAlaIleIleGlyLysGlnGlyGlnHisIleLysGlnLeuSerArgPheAla 435
QY 1441 GGAGCCTCTATCAAGATTGCCCTCGCGAAGGCCAGACGTCAGCGAAAGAGATGGTATC 1500
```

```
Db 436 GlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLysValArgMetValIle 455
QY 1501 ATCACCGGCCACCGAAGCCCGAGTTCAAGGCCCGAGGACGGATCTTTGGGAAACTGAAA 1560
Db 456 IleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysIleLys 475
QY 1561 GAGGAAACTTCTTTAACCCCAAGAAAGTGAAGCTGGAAGCGCATATCAGAGTGCCC 1620
Db 476 GluGluAsnPheValSerProLysGluGluValLysLeuGluAlaHisIleArgValPro 495
QY 1621 TCTTCCACAGCTGCGCGGTGATTGCCAAAGGTGGCAAGACCGTGAACGAACCTGCAGAAC 1680
Db 496 SerPheAlaAlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsn 515
QY 1681 TTAACCAAGTGCAGAACTCATCGTCCCTCGTGACCAAAACGCCAGATGAAATGAGGAAGT 1740
Db 516 LeuSerSerAlaGluValValProArgAspGlnThrProAspGluAsnAspGlnVal 535
QY 1741 ATCGTCAGAAATTATCGGGCACTTCTTTGCTAGCCAGACTGCACAGCGCAAGATCAGGAA 1800
Db 536 ValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGlnArgLysIleGlnGlu 555
QY 1801 ATTGTCAACAGCTGAAGCAGCAGCAGCAGAAA 1833
Db 556 IleLeuThrGlnValLysGlnHisGlnGlnGln 566

RESULT 9
ABB75053
ID ABB75053 standard; protein; 579 AA.
XX
AC ABB75053;
XX
DT 01-MAY-2002 (first entry)
XX
DE Human lung tumour L523S recombinant protein sequence SEQ ID NO:446.
XX
KW Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;
KW immune response.
XX
OS Homo sapiens.
XX
PN WO200200174-A2.
XX
PD 03-JAN-2002.
XX
PF 28-JUN-2001; 2001WO-US021065.
XX
PR 28-JUN-2000; 2000US-00606421.
PR 02-AUG-2000; 2000US-00630940.
PR 21-AUG-2000; 2000US-00643597.
PR 15-SEP-2000; 2000US-00662786.
PR 09-OCT-2000; 2000US-00685696.
PR 12-DEC-2000; 2000US-00735705.
PR 07-MAY-2001; 2001US-00850716.
XX
PA (CORI-) CORIXA CORP.
XX
PI Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
PI McNeill PD, Fanger N, Retter MW, Marnierakis M, Fanger GR;
PI Vedvick TS, Carter D, Watanabe Y, Peckham DW;
XX
DR WPI; 2002-090513/12.
DR N-PSDB; ABL49297.
XX
PT Polynucleotides encoding lung tumor polypeptides, useful for treating
PT lung cancer or stimulating an immune response.
XX
PS Claim 2; Page 365-367; 374pp; English.
XX
CC The present invention describes human lung tumour proteins. Human lung
CC tumour proteins and polynucleotides have cytostatic and immunostimulant
CC activities, and can be used in vaccine production. Compositions
```

CC comprising the lung tumour proteins, polynucleotides, antibodies, fusion
 CC proteins, T cell populations, or antigen presenting cells that express
 CC the lung tumour proteins are useful for treating lung cancer or
 CC stimulating an immune response. ABL48959 to ABL49300 and ABB74946 to
 CC ABB75070 represent sequences used in the exemplification of the present
 CC invention
 XX
 SQ Sequence 579 AA;

Alignment Scores:

Pred. No.: 6,18e-168 Length: 579
 Score: 1957.50 Matches: 387
 Percent Similarity: 79.19% Conservative: 81
 Best Local Similarity: 65.48% Mismatches: 94
 Query Match: 31.98% Indels: 29
 DB: 5 Gaps: 8

US-09-270-437D-6 (1-3412) x ABB75053 (1-579)

QY	73	ATGAACAAGCTTTATCATCGGAACCTGAGCCCGCGCTCAGCCGCGAGCACTCCGCGCAG	132
DB	1	MetAsnLysLeuTyrIleGlyAsnLeuSerGluAsnAlaAlaProSerAspLeuGluSer	20
QY	133	CTCTTTGGGGACAGAGCTGCCCTCGGGGACAGAGTCTCTGTAAGTCCGGCTACGCC	192
DB	21	IlePheLysAspAlaLysIleProValSerGlyProPheLeuValLysThrGlyTyrAla	40
QY	193	TTGCTGGAGTACCCCGACAGAACTGGCCCATCGCGCCATCGAGACCTCTCGGGTAA	252
DB	41	PheValAspCysProAspGluSerIlePheValLeuLysAlaIleGluAlaLeuSerGlyLys	60
QY	253	GTGGAATTGCTGGGAAATCATGGAAGTTGATTACTCAGTCTCTTAAAGCTAAGGAGC	312
DB	61	IleGluLeuHisGlyLysProIleGluValGluHisSerValProLysArgGlnArgIle	80
QY	313	AGGAAATTGATTCGAAACATCCCTCTCAGTGCAGTGGAGGTTTGGATGGAATT	372
DB	81	ArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspSerLeu	100
QY	373	TTGGCTCAATATGGACAGTGGAGAAATGTGGAAGTCAACAGAGCAACAGCAACCGCC	432
DB	101	LeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAspSerGluThrAla	120
QY	433	GTGTCACAGTCATATGCAACAGAGAGAAAGCAAAATAGCCATGGAAGCTAAGC	492
DB	121	ValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn	140
QY	493	GGGCATCAGTTTGAACTACTCTTCAAGATTTCCTACATCCCGGATGAAGAGTGAGC	552
DB	141	GlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAspGluThrAlaAla	160
QY	553	TCCCTTTCGCCCCCTCAG-----CGAGCCAGCGT-----GGGACCACTCTTCCCGG	600
DB	161	GlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlyGlnArgGlySerSer	180
QY	601	GAGCAAGCCAGCCCTCGGGGCACTTCTCAGCCACAGACAGATTGATTCCTCCGCTGGG	660
DB	181	ArgGlnGly---SerProGlySerValSerLysGlnLysProCysAspLeuProLeuArg	199
QY	661	ATCTGTGTCACCCAGTTTGTGTCATCATCGGAAGAGGAGGCTTACCATTAAG	720
DB	200	LeuLeuValProThrGlnPheValGlyAlaIleIleGlyLysGluGlyAlaThrIleArg	219
QY	721	AACATCACTAAGCAGACCCAGTCCCGGTAGATATCCATGAAAGAGAACTCTGGAGCT	780
DB	220	AsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGlyAla	239
QY	781	GCAGAGAGCTGTCACTCATTCATGCCCCAGAGGGAATTCTGAGCATCGCGCATG	840
DB	240	AlaGluLysSerIleThrIleLeuSerThrProGluGlyThrSerAlaAlaCysLysSer	259
QY	841	ATTCTTGAATCATGCAGAGAGCGAGATGAGACCAACTAGCCGAGAGATTCCTCTG	900

DB	260	IleLeuGluIleMetHisLysGluAlaGlnAspIleLysPheThrGluGluIleProLeu	279
QY	901	AAATCTTGGCACACAATGGCTTGGTAAGACTGATTGAAAGAGAGCGAGAATTG	960
DB	280	LysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsnLeu	299
QY	961	AAGAAATTTGAACATGAAACAGGACCAAGATAACAATCTCATCTTTCAGGATTGAGC	1020
DB	300	LysLysIleGluGlnAspThrAspThrLysIleThrIleSerProLeuGlnGluLeuThr	319
QY	1021	ATATACAAACCCGAAAGAACCATCATCTGTGAAGGCACAGTTGAGCCCTGTGCCAGTGC	1080
DB	320	LeuTyrAsnProGluArgThrIleThrValLysGlyAsnValGluThrCysAlaLysAla	339
QY	1081	GAGATAGAGATTATGAAGAGCTGGCTGAGCCCTTTGAAATGATATGCTGGCTTTAAC	1140
DB	340	GluGluGluIleMetLysLysIleArgGluSerTyrGluAsnAspIleAlaSerMetAsn	359
QY	1141	CAACAAGCCAATCTGATCCCAAGGTTGAACCTCAGCGCACCTTGGCATCTTTTCAACAGA	1200
DB	360	LeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeuGlyLeuPhe-----	376
QY	1201	CTGTCCGTGCTATCTCACACAGCAGGCGCCCGCGAGCTCCCGCGTGCCTCCCTTACCAC	1260
DB	377	-----ProProThrSer-----GlyMetProProProThr-----	386
QY	1261	CCCTTCACTACCCACTCCGGATACTTCTCCAGCCTGTACCCCATCACCAGTTTGGCCCG	1320
DB	387	-----SerGlyProProSerAlaMetThrPro-----Pro	396
QY	1321	TTCCCGCATCATCTCTTATCCAGACGAGAGATTGTAATCTTCTTTCATCCCAACCCAG	1380
DB	397	TyrProGlnPheGluGln---SerGluThrGluThrValHisLeuPheIleProAlaLeu	415
QY	1381	GCTGTGGCGCCATCATCGGAAAGAGGGGSCACACATCAACAGCTGCGGAGATTCGCC	1440
DB	416	SerValGlyAlaIleIleGlyLysGlnGlyGlnHisIleLysGlnLeuSerArgPheAla	435
QY	1441	GGAGCTCTATCAGATTGCCCTCGGAGAGGCCAGACCTCAGCGAAAGGATGGTCAATC	1500
DB	436	GlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLysValArgMetValIle	455
QY	1501	ATCACCGGCGCCAGGAGCCAGCTTCAAGCCCGAGGACCGATCTTTGGAAACTGAAA	1560
DB	456	IleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysIleLys	475
QY	1561	GAGAAATCTTCTTTAAACCCCAAGAGAGTGAAGCTGGAGCGCATATCAGAGTGCC	1620
DB	476	GluGluAsnPheValSerProLysGluGluValLysLeuGluAlaHisIleArgValPro	495
QY	1621	TCTTCCACAGCTGCGCGGTGATTGGCAAGGTGGCAAGCCGTAAGCACTGACAGAC	1680
DB	496	SerPheAlaAlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsn	515
QY	1681	TTAACCAAGTCAGAGTCAATCGTCCCTCGTCAACCAAGCCAGATGAAATGAGAAAGTG	1740
DB	516	LeuSerSerAlaGluValValProArgAspGlnThrProAspGluAsnAspGlnVal	535
QY	1741	ATCGTCAGATTATCGGCGACTTCTTGTAGCCAGACTGCACAGCCAGATCAGGAA	1800
DB	536	ValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGlnArgLysIleGlnGlu	555
QY	1801	ATTGTACACAGAGTGAAGCAGCAGCAGAGAAA	1833
DB	556	IleLeuThrGlnValLysGlnHisGlnGlnGln	566

RESULT 10
 ABB74997
 ID ABB74997 standard; protein; 579 AA.
 XX
 AC ABB74997;
 XX 01-MAY-2002 (first entry)
 DT

```

XX DE Human lung tumour L523S protein sequence SEQ ID NO:348.
XX KW Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine;
XX KW immune response.
XX OS Homo sapiens.
XX XX
XX PN WO200200174-A2.
XX XX
XX XX 03-JAN-2002.
XX XX
XX XX 28-JUN-2001; 2001WO-US021065.
XX XX
XX PR 28-JUN-2000; 2000US-00606421.
XX PR 02-AUG-2000; 2000US-00630940.
XX PR 21-AUG-2000; 2000US-00643597.
XX PR 15-SEP-2000; 2000US-00662786.
XX PR 09-OCT-2000; 2000US-00685696.
XX PR 12-DEC-2000; 2000US-00735705.
XX PR 07-MAY-2001; 2001US-00850716.
XX XX
XX PA (CORI-) CORIXA CORP.
XX XX
XX PI Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
XX PI McNeill PD, Fanger N, Retter MW, Warnerakis M, Fanger GR;
XX PI Vedvick TS, Carter D, Watanabe Y, Peckham DW;
XX XX
XX DR WPI; 2002-090513/12.
XX DR N-PSDB; ABL49254.
XX XX
XX PT Polynucleotides encoding lung tumor polypeptides, useful for treating
XX PT lung cancer or stimulating an immune response.
XX PS
XX PS Example 2; Page 330-332; 374pp; English.
XX XX
XX CC The present invention describes human lung tumour proteins. Human lung
XX CC tumour proteins and polynucleotides have cytostatic and immunostimulant
XX CC activities, and can be used in vaccine production. Compositions
XX CC comprising the lung tumour proteins, polynucleotides, antibodies, fusion
XX CC proteins, T cell populations, or antigen presenting cells that express
XX CC the lung tumour proteins are useful for treating lung cancer or
XX CC stimulating an immune response. ABL48959 to ABL49300 and ABB74946 to
XX CC ABB75070 represent sequences used in the exemplification of the present
XX CC invention
XX XX
XX SQ Sequence 579 AA;

Alignment Scores:
Pred. No.: 6.18e-168 Length: 579
Score: 1957.50 Matches: 387
Percent Similarity: 79.19% Conservative: 81
Best Local Similarity: 65.48% Mismatches: 94
Query Match: 31.98% Indels: 29
DB: 5 Gaps: 8

US-09-270-437D-6 (1-3412) x ABB74997 (1-579)

QY 73 ATGAACAAAGCTTTACATCGGGAACTGAGCCCGCCCGCTACCGCCGACGACCTCCGGCAG 132
Db 1 MetAsnLysLeuTyrlleGlyAsnLeuSerGluAsnAlaProSerAspLeuGluSer 20
QY 133 CTCTTTGGGACAGAGAGCTGCCCTCGCGGGACAGGTCCTGCTGAAGTCGGGTAGCC 192
Db 21 IlePheLysAspAlaLysIleProValSerGlyProPheLeuValLysThrGlyTyra 40
QY 193 TTCGTGGACTACCCGACCACTGGGCGCATCGCGCCATCGAGACCCTCTCGGTA 252
Db 41 PheValAspCysProAspGluSerTrpAlaLeuLysAlaIleGluAlaLeuSerGlyLys 60
QY 253 GTGGAATTGATGGGAAATCATGAAGTTGATTACTACTAGCTCTATAAAGACTAAGGAGC 312
Db 61 IleGluLeuHisGlyLysProIleGluValGluHisSerValProLysArgGlnArgIle 80

313 AGGAAAATTGAGATTGGAACATCCCTCCTCAGTCAGTGGAGGCTGTGGTGGACTT 372
81 ArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspSerLeu 100
373 TTGGCTCAATATGGGACAGTGGAGAATGTGGAACAAGTCAACACAGACACAGAAACCGCC 432
101 LeuValGlnTyrglyValValGluSerCysGluGlnValAsnThrAspSerGluThrAla 120
433 GTTGTCAACGTCAATATGCAACAGAGAAAGCAAAATAGCCATGGAGAGCTTAAGC 492
121 ValValAsnValThrTyrsSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn 140
493 GGGCATCAGTTTGGAGAACTACTCTTCAAGATTTCCTACATCCCGATCGAGAGGTGAGC 552
141 GlyPheGlnLeuGlnAsnPheThrLeuLysValAlaIleProAspGluThrAlaAla 160
553 TCCCTTCGCCCTCCCTCAG-----CGAGCCACAGCGT-----GGGACCACTCTTCCCGG 600
161 GlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlyGlnArgLysSerSer 180
601 GAGCAAGCCACGCCCTCGGGGCACTTCTCAGCCAGACAGATGATTGATTTCCGCTGCGG 660
181 ArgGlnGly---SerProGlySerValSerLysGlnLysProCysAspLeuProLeuArg 199
661 ATCCTGCTCCACCCAGTTTGTGTCCTCATCGGATCGGAAAGGAGGCTTGACATAAAG 720
200 LeuLeuValProThrGlnPheValGlyAlaIleGlyLysGluGlyAlaThrIleArg 219
721 AACATCACTAAGCAGACCCAGTCCCGGTAGATATCATAGAAAAGAGAACTCTGGAGCT 780
220 AsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGlyAla 239
781 GCAGAGAAGCTGTCCACATCCATGCCATGCCACCCAGAGGGGACTTCTGAAGATCCCGCATG 840
240 AlaGluLysSerIleThrIleLeuSerThrProGluGlyThrSerAlaAlaCysLysSer 259
841 ATTCTGTAATCATGCAAGAGGCGAGATGAGACCACTAGCCGAGAGATTCCTCTG 900
260 IleLeuGluIleMetHisLysGluAlaGlnAspIleLysPheThrGluGluIleProLeu 279
901 AAAATCTGGCACACAATGGCTTGGTGAAGACTGATTGAAAAGAGGAGCAGAAATTG 960
280 LysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsnLeu 299
961 AAGAAAATTGAACATGAACAGGACCAAGATACCAATCTCATCTTTTCAGAGATTTCAGC 1020
300 LysLysIleGluGlnAspThrAspThrLysIleThrIleSerProLeuGlnGluLeuThr 319
1021 ATATACAAACCGGAAGAACCATCATCTGTGAAGGGCACAGTTCAGGCCTGTCGAGTCT 1080
320 LeuTyrsAsnProGluArgThrIleThrValLysGlyAsnValGluThrCysAlaLysAla 339
1081 GAGATAGAGATTATGAAGAGCTGCGTGAAGCCCTTTGAAAATGATATGCTGGCTGTTAAC 1140
340 GluGluGluIleMetLysLysIleArgGluSerTyrgluAsnAspIleAlaSerMetAsn 359
1141 CAACAGCCCAATCTGATCCAGGGTTGAACCTCAGCGCACTTGGCATCTTTTCAACAGGA 1200
360 LeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeuGlyLeuPhe----- 376
1201 CTGTCGCTGTATCTCCACAGGAGCCCGCGGAGCTCCCGCGCTCCCGCTGCGCCCTACCAC 1260
377 -----ProProThrSer-----GlyMetProProProThr----- 386
1261 CCCTTCACTACCACTCCGGATACTTCTCCAGCTGTACCCCATCACCAGTTTGGCCCG 1320
387 -----SerGlyProProSerAlaMetThrPro-----Pro 396
1321 TTCCCGCATCATCACTCTTATCCAGAGCAGGAGATTGTGAATCTCTTCATCCCAACCCAG 1380
397 TyrProGlnPheGluGln---SerGluThrGluThrValHisLeuPheIleProAlaLeu 415

```


Db 240 AladgluysSerIleThrIleuSerThrProGluGlyThrSerAlaAlaCysLysSer 259
 Qy 841 ATTCTTGAATCATGACAGAGCAGATGAGACCAAACTAGCGAAGAGATTCCTGTG 900
 Db 260 ILeuGluIleMetHisLysGluAlaGlnAspIleLysPheThrGluGluIleProLeu 279
 Qy 901 AAAATCTTGGCACACAATGGCTGTGTTGGAGACTGATTGGAAGAGAGGAGCAAAATTG 960
 Db 280 LysIleLeuAlaHisAsnAsnPheValGlyArgIleGlyLysGluGlyArgAsnLeu 299
 Qy 961 AAGAAAATTGAACATGAACAGGACCAAGATATAACAATCTCTTTTCGACGATTGAGC 1020
 Db 300 LysLysIleGluGlnAspThrAspThrLysIleThrIleSerProLeuGlnGluLeuThr 319
 Qy 1021 ATATACACCCGGAAGAACCATCTCTGTGAAGGCACAGTGTGAGGCTGTGCCAGTCT 1080
 Db 320 LeuTyrAsnProGluArgThrIleThrValLysGlyAsnValGluThrCysAlaLysAla 339
 Qy 1081 GAGATAGAGATTATGAAGAAGCTCGTGCAGGCTTGAACCTCAGCGCACTTGGCATCTTTTCAACAGGA 1200
 Db 360 LeuGlnAlaHisLeuIleProGlyLeuAsnLeuAlaLeuGlyLeuPhe----- 376
 Qy 1201 CTGTCGGTGCTATCTCCACGACGAGGCGCCGCGGAGCTCCCGCGTGC-----TACCAC 1260
 Db 377 -----ProThrSer-----GlyMetProProThr----- 386
 Qy 1261 CCCTTCACTACCCACTCGGATCTTCTCCAGCTGTACCCCATCACCATTGGCCCG 1320
 Db 387 -----SerGlyProProSerAlaMetThrPro-----Pro 396
 Qy 1321 TTCCCGCATCATCTCTTATCCAGCAGCAGAGATTGTGAATCTCTTTCATCCCAACCCAG 1380
 Db 397 TyrProGlnPheGluGln---SerGluThrGluThrValHisLeuPheIleProAlaLeu 415
 Qy 1381 GCTGTGGCGGCATCATCGGGAAGAGAGGGGCAACATCAACAGCTGGCGAGATTGCGC 1440
 Db 416 SerValGlyAlaIleIleGlyLysGlnGlyGlnHisIleLysGlnLeuSerArgPheAla 435
 Qy 1441 GGAGCTCTATCAGATTGCTTCCGCGAGGCGCCAGCGTCAGCGAAGAGATGTCATC 1500
 Db 436 GlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLysValArgMetValIle 455
 Qy 1501 ATCACCGGCGCCAGCGAGCCAGTTCAAGCGCCAGGACGATCTTTGGGAAACTGAAA 1560
 Db 456 IleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysIleLys 475
 Qy 1561 GAGAAAACCTTCTTTAACCCCAAGAGAACTGAGCTGGAAGCGCATATCAGAGTGC 1620
 Db 476 GluGluAsnPheValSerProLysGluGluValLysLeuGluAlaHisIleArgValPro 495
 Qy 1621 TCTTCCAGCTGCGCGGTGATTGGCAAGAGTGCAGACCGTGAACGACTCGAGAC 1680
 Db 496 SerPheAlaIaGlyArgValIleGlyLysGlyThrValAsnGluLeuGlnAsn 515
 Qy 1681 TTAACCAAGTCAGAGATCATCGTCCCTGTGACCAAAAGCCAGATGAAATAGAGAGTG 1740
 Db 516 LeuSerSerAlaGluValValProArgAspGlnThrProAspGluAsnAspGlnVal 535
 Qy 1741 ATCGTCAGANTATCGGGCACTCTTCTGTCAGCCAGCTGACAGCGCAAGATCAGGAA 1800
 Db 536 ValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGlnArgLysIleGlnGlu 555
 Qy 1801 ATTGTACAAAGGTGAAGCACAGGAGCAGAAA 1833
 Db 556 IleLeuThrGlnValLysGlnHisGlnGlnGln 566

RESULT 12
 ABP61917

ID XX ABP61917 standard; protein; 579 AA.
 AC XX ABP61917;
 DT XX 07-OCT-2002 (first entry)
 XX Human lung cancer associated protein sequence SEQ ID NO:348.
 XX Human; lung cancer; lung tumour; cytostatic; gene therapy; vaccine.
 OS Homo sapiens.
 XX WO200247534-A2.
 XX 20-JUN-2002.
 XX 30-NOV-2001; 2001WO-US047576.
 XX 12-DEC-2000; 2000US-00735705.
 XX 07-MAY-2001; 2001US-00850716.
 XX 28-JUN-2001; 2001US-00897778.
 PA (CORI-) CORIXA CORP.
 XX Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
 PI McNeill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedwick TS;
 PI Carter D, Watanabe Y, Peckham DW, Cai F, Foy TM;
 DR WPI; 2002-583465/62.
 DR N-PSDB; ABQ92440.
 XX Novel lung carcinoma polynucleotide sequences and polypeptides encoded by
 PT the polynucleotides, useful in pharmaceutical compositions such as
 PT vaccines and as markers to indicate the presence of lung cancer.
 XX Example 2; Page 337-339; 381pp; English.
 CC The present invention describes isolated human lung carcinoma
 CC polynucleotides (I) and polypeptides (II). (I) and (II) have cytostatic
 CC activity, and can be used in gene therapy and in vaccines. Compositions
 CC comprising (I) or (II) can be used for stimulating an immune response in
 CC a patient and for treating lung cancer in a patient. Oligonucleotides of
 CC (I) can be used for detecting the presence of a cancer in a patient, by
 CC obtaining a biological sample from the patient, contacting the biological
 CC sample with the oligonucleotide, detecting in the sample, an amount of
 CC polynucleotide that hybridises to the oligonucleotide and comparing the
 CC amount of polynucleotide that hybridises to the oligonucleotide to a
 CC predetermined cut-off value, and determining the presence of a cancer in
 CC the patient. (I) and (II) are useful in pharmaceutical compositions, e.g.
 CC vaccines. (I) is useful as a marker to indicate the presence or absence
 CC of a cancer such as lung cancer. ABQ92145 to ABQ92486 and ABP61866 to
 CC ABP61992 represent sequences used in the exemplification of the present
 CC invention

SQ Sequence 579 AA;

Alignment Scores:
 Pred. No.: 6.18e-168 Length: 579
 Score: 1957.50 Matches: 387
 Percent Similarity: 79.19% Conservative: 81
 Best Local Similarity: 65.48% Mismatches: 94
 Query Match: 31.98% Indels: 29
 DB: 5 Gaps: 8

US-09-270-437D-6 (1-3412) x ABP61917 (1-579)

Qy 73 ATGACAGCTTTACATCGGAGACCTGAGCCCGCGTACCGCCGACGACCTCCGCGAG 132
 Db 1 MetAsnLysLeuTyrIleGlyAsnLeuSerGluAsnAlaAlaProSerAspLeuGluSer 20
 Qy 133 CTCTTTGGGACAGGAGCTGCCCTGGCGGACAGGTCTCTGCTGAGTCCGCGCTAGCC 192
 Db 21 IlePheLysAspAlaLysIleProValSerGlyProPheLeuValLysThrGlyTyrAla 40

QY	193	TTCTGGGACTACCCGACAGAACTGGGCATCCGCGCCATCGAGACCTCTCGGTAAA	252
Db	41	PheValAspCysProAspGluSerThrPalaLeuLysAlaIleGluAlaLeuSerGlyLys	60
QY	253	GTGGAATTGCATGGGAAATCATGGAAGTTGATTACTCAGTCTCTAAAAGCTAAGGAGC	312
Db	61	IleGluLeuHisGlyLysProIleGluValGluHisSerValProLysArgGlnArgIle	80
QY	313	AGGAAATTCAGATTGCAAAACATCCCTCTCCTACCTGCAGTGGAGGTGTGATGACTT	372
Db	81	ArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspSerLeu	100
QY	373	TTGGCTCAATATGGGACAGTGGAGAATGGGAACAAGTCAACACAGACACAGAAACCGCC	432
Db	101	LeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAspSerGluThrAla	120
QY	433	GTGTCAACGTCACATATCAACAAGAGAAAGCAAAATAGCCATGCGAGAGCTAAGC	492
Db	121	ValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn	140
QY	493	GGGCATCATGTTGAGAACTACTCTTCAGATTCTACATCCCGATCCGAGGTGAGC	552
Db	141	GlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAspGluThrAlaAla	160
QY	553	TCCCTCTCCGCCCTCAG-----CGAGCCACGCGT-----GGGGACCACTCTCCCGG	600
Db	161	GlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlyGlnArgGlySerSer	180
QY	601	GAGCAAGGCCACGCCCTCGGGGCATCTCTCAGGCCAGACAGATTGATTCGCCGTGGG	660
Db	181	ArgGlnGly---SerProGlySerValSerLysGlnLysProCysAspLeuProLeuArg	199
QY	661	ATCTGTGTCGCCACCCAGTTTGTGTGTCATCATCGGAAGAGGGCTTGACCATAAAG	720
Db	200	LeuLeuValProThrGlnPheValGlyAlaIleIleGlyLysGluGlyAlaThrIleArg	219
QY	721	AACATCATCAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAGAGAACTCTCGAGCT	780
Db	220	AsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGlyAla	239
QY	781	GCAGAGACCTGTACCATCCATGCCATGCCACCCAGAGGGACATTCTGAACATGCCGATG	840
Db	240	AlaGluLysSerIleThrIleLeuSerThrProGluGlyThrSerAlaAlaCysLysSer	259
QY	841	ATTCTTGAATCATGCAGAGAGGCAGATGAGACCAACTAGCCGACAGATTCTCTCG	900
Db	260	IleLeuGluIleMethHisLysGluAlaGlnAspIleLysPheThrGluGluIleProLeu	279
QY	901	AAAACTTGGCACACAATGGCTTGGTTGGAAGACTGATTGGAAAAGAGCCAGAAATTG	960
Db	280	LysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsnLeu	299
QY	961	AGAAATTTGAATGAACAGGACCAAGATAACAAATCTCATCTTTTGGAGATTGAGC	1020
Db	300	LysLysIleGluGlnAspThrAspThrLysIleThrIleSerProLeuGlnGluLeuThr	319
QY	1021	ATATACACCCGAAAGACCATCACTGTGAAGGGCAGAGTTGAGCCTGTGCCAGTGCT	1080
Db	320	LeuTyrAsnProGluArgThrIleThrValLysGlyAsnValGluThrCysAlaLysAla	339
QY	1081	GAGATAGATTATGAAGAGCTGCGTGAAGGCTTTGAAATCATATGCTGCTGTAAAC	1140
Db	340	GluGluGluIleMethLysLysIleArgGluSerTyrGluAsnAspIleAlaSerMetAsn	359
QY	1141	CAACAGCAATCTGATCCAGGGTTGAACCTCAGCGGCACTTGGCATCTTTTCAACAGA	1200
Db	360	LeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeuGlyLeuPhe-----	376
QY	1201	CTGTCCGTGTATCTCACCAGCAGGGCCCCGCGAGCTCCCGCTGCCCCCTACCAC	1260
Db	377	-----ProProThrSer-----GlyMetProProThr-----	386
QY	1261	CCCTTCACTACCCACTCCGGATACCTTCTCCAGCTGTATCCCCCATCACACAGTTTGGCCGG	1320
Db	387	-----SerGlyProProSerAlaMetThrPro-----Pro	396
QY	1321	TTCCGCGATCATCATCTTTATCCAGAGCAGAGATTGGAATCTTCTATCCCAACCCAG	1380
Db	397	TyrProGlnPheGluGln---SerGluThrGluThrValHisLeuPheIleProAlaLeu	415
QY	1381	GCTGTGGCGCCATCATCGGAAGAAGGGGACACATCAACACAGCTGCGAGATTCCGC	1440
Db	416	SerValGlyAlaIleIleGlyLysGlnGlnHisIleLysGlnLeuSerArgPheAla	435
QY	1441	GGAGCTCTATCAAGATTGCCCTCGGAAGCCAGACGTCAGCGAAAGAGTGTCTATC	1500
Db	436	GlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLysValArgMetValIle	455
QY	1501	ATCACCGGCGCACCCGGAAGCCAGTTCAAGGCCAGGACCGGATCTTTGGGAAACTGAAA	1560
Db	456	IleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysIleLys	475
QY	1561	GAGGAAAATCTCTTTAACCCCAAGAAAGTGAAGCTGGAAGCGCATATCAGAGTGCC	1620
Db	476	GluGluAsnPheValSerProLysGluGluValLysLeuGluAlaHisIleArgValPro	495
QY	1621	TCTTCCACAGCTGGCGGGTGTATTGGCAAGAGTGGCAAGACCGTGAAACCACTGCAGAAC	1680
Db	496	SerPheAlaAlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsn	515
QY	1681	TTAACCACTGCAGAAAGTATCGTGTCTGTGACCAAAACGCCAGATGAAAATCAGGAAGTG	1740
Db	516	LeuSerSerAlaGluValValProArgAspGlnThrProAspGluAsnAspGlnVal	535
QY	1741	ATCGTCAGAAATTATCGGCACCTTCTTGTCTGCCAGACTGCACAGCGCAAGATCAGGGAA	1800
Db	536	ValValLysIleThrGlyHisPheLysAlaCysGlnValAlaGlnArgLysIleGlnGlu	555
QY	1801	ATTGTACAAACAGGTGAAGCAGCAGGACGAGAAA	1833
Db	556	IleLeuThrGlnValLysGlnHisGlnGlnGln	566
RESULT 13			
ABP61974			
ID	ABP61974	standard; protein; 579 AA.	
XX	XX	ABP61974;	
AC	XX	ABP61974;	
XX	XX	07-OCT-2002 (first entry)	
DT	XX	Human lung cancer associated protein sequence SEQ ID NO:449.	
DE	XX	Human; lung cancer; lung tumour; cytostatic; gene therapy; vaccine.	
KW	XX	Homo sapiens.	
OS	XX	WO200247534-A2.	
XX	XX	20-JUN-2002.	
PN	XX	30-NOV-2001; 2001WO-US047576.	
PD	XX	12-DEC-2000; 2000US-00735705.	
PF	XX	07-MAY-2001; 2001US-00850716.	
XX	XX	28-JUN-2001; 2001US-00897778.	
XX	XX	(CORI-) CORIXA CORP.	
PA	XX	Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;	
XX	XX	McNeill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedvick TS;	
PI	XX	Carter D, Watanabe Y, Peckham DW, Cai F, Foy TW;	
XX	XX	WPI: 2002-583465/62.	
DR	XX	N-PSDB; ABQ92485.	
DR	XX		

PT Novel lung carcinoma polynucleotide sequences and polypeptides encoded by
 PT the polynucleotides, useful in pharmaceutical compositions such as
 PT vaccines and as markers to indicate the presence of lung cancer.

XX PS Claim 9; Page 375-377; 381pp; English.

XX CC The present invention describes isolated human lung carcinoma
 CC polynucleotides (I) and polypeptides (II). (I) and (II) have cytostatic
 CC activity, and can be used in gene therapy and in vaccines. Compositions
 CC comprising (I) or (II) can be used for stimulating an immune response in
 CC a patient and for treating lung cancer in a patient. Oligonucleotides of
 CC (I) can be used for detecting the presence of a cancer in a patient, by
 CC obtaining a biological sample from the patient, contacting the biological
 CC sample with the oligonucleotide, detecting in the sample, an amount of
 CC polynucleotide that hybridises to the oligonucleotide and comparing the
 CC amount of polynucleotide that hybridises to the oligonucleotide to a
 CC predetermined cut-off value, and determining the presence of a cancer in
 CC the patient. (I) and (II) are useful in pharmaceutical compositions, e.g.
 CC vaccines. (I) is useful as a marker to indicate the presence or absence
 CC of a cancer such as lung cancer. ABQ2145 to ABQ2486 and ABP61866 to
 CC ABP61992 represent sequences used in the exemplification of the present
 CC invention

XX SQ Sequence 579 AA;

Alignment Scores:

Pred. No.: 6.18e-168 Length: 579
 Score: 1957.50 Matches: 387
 Percent Similarity: 79.1% Conservative: 81
 Best Local Similarity: 65.48% Mismatches: 94
 Query Match: 31.98% Indels: 29
 DB: 5 Gaps: 8

US-09-270-437D-6 (1-3412) x ABP61974 (1-579)

QY 73 ATGAACAGCTTTATACGCGGAACCTGAGCCCGCCCGCTCACCGCGAGCAGCTCCGGCAG 132
 DB 1 MetAsnLysLeuTyrlleGlyAsnLeuSerGluAsnAlaalaProSerAspLeuGluSer 20
 QY 133 CTCTTTGGGACAGGAGCTGCCCTCGCGGAGCAGTCTCTGTGAAGTCGGGTACGCC 192
 DB 21 IlePheLysAspAlaLysIleProValSerGlyProPheLeuValLysThrGlyTyra 40
 QY 193 TTCTGTGACTACCCGACAGAACTGGCCATCCGCGCCATCGAGCCCTCTCGGTAAA 252
 DB 41 PheValAspCysProAspGluSerTrpAlaLeuLysAlaIleGluAlaLeuSerGlyLys 60
 QY 253 GTGGAATTGCGGAAATCATGGAAGTTGATTACTCAGTCTCTAAAGAGCTAAGGAGC 312
 DB 61 IleGluLeuHisGlyLysProIleGluValGluHisSerValProLysArgGlnArgIle 80
 QY 313 AGGAAATTCAGATTGGAACATCCCTCTACCTGAGTGGGAGGTGTGGATGGACTT 372
 DB 81 ArgLysLeuGlnIleArgAsnIleProHisLeuGlnTrpGluValLeuAspSerLeu 100
 QY 373 TTGSCCTCAATATGGACAGTGGAGAACTGGAACAAGTCAACAGTCAACAGACAGAACCGCC 432
 DB 101 LeuValGlnTrpGlyValValGluSerCysGluGlnValAsnThrAspSerGluThrAla 120
 QY 433 GTTCTCAAGCTCATATGCAACAGAGAGAAAGCAAAATAGCCATGGAGAGCTAAGC 492
 DB 121 ValValAsnValThrTyrrSerSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn 140
 QY 493 GGGCATCAGTTTGAGAACTACTCTTCAAGATTTCCTACATCCCGGATGAAGAGGTGAGC 552
 DB 141 GlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrlleProAspGluThrAlaAla 160
 QY 553 TCCCTTCGCCCTCCAG-----CGAGCCAGAGT-----GGGACCACCTCTCCGG 600
 DB 161 GlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlyGlnArgGlySerSer 180
 QY 601 GAGCAAGGCCACGCCCTCGGGGACCTTCTCAGGCCAGACAGATTGATTTCCTCCGCTCGG 660
 DB 516 LeuSerSerAlaGluValValProArgAspGlnThrProAspGluAsnAspGlnVal 535

DB 181 ArgGlnGly---SerProGlySerValSerLysGlnLysProCysAspLeuProLeuArg 199
 QY 661 ATCCTGTCCCCACCCAGTTTGTGTGCCATCATCGAAAGAGGGCGCTTGACCATAAAG 720
 DB 200 LeuLeuValProThrGlnPheValGlyAlaIleGlyLysGluGlyAlaThrIleArg 219
 QY 721 AACATCACTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAAGAACTCTGGAGCT 780
 DB 220 AsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGlyAla 239
 QY 781 GCAGAGAAGCTGTGCACATCCATCCACCCAGAGGGGACTTCTGAAGCATGCCGATG 840
 DB 240 AlaGluLysSerIleThrIleLeuSerThrProGluGlyThrSerAlaAlaCysLysSer 259
 QY 841 ATTCTTGAATCATGCGAAGAGGCGAGATGAGACCAAACTAGCCGAGAGATTCTCTGTG 900
 DB 260 IleLeuGluIleMetHisLysGluAlaGlnAspIleLysPheThrGluGluIleProLeu 279
 QY 901 ABAATCTTGGCACACATGCTGTGTGGAGACTGATTGGAAAAGAGAGCAGAAATTG 960
 DB 280 LysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsnLeu 299
 QY 961 AAGAAATTCGAATGAACAGAGGACCAAGATACAAATCTCATCTTTCAGGATTTCAGC 1020
 DB 300 LysLysIleGluGlnAspThrAspThrLysIleThrIleSerProLeuGlnGluLeuThr 319
 QY 1021 ATATACAACCCGGAAGAACCATCATCTGTGAAGGGCACAGTTGAGGCTGTGCCAGTCT 1080
 DB 320 LeuTyrrAsnProGluArgThrIleThrValLysGlyAsnValGluThrCysAlaLysAla 339
 QY 1081 GAGATAGAGATTATGAAGAGCTGCGTGAGCGCTTTGAAAATGATATGCTGGCTGTAAAC 1140
 DB 340 GluGluGluIleMetLysLysIleArgGluSerTyrrGluAsnAspIleAlaSerMetAsn 359
 QY 1141 CAACAACCAATCTGATCCAGGGTGAACCTCAGCGCACTTGCACTCTTTCACACAGGA 1200
 DB 360 LeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeuGlyLeuPhe----- 376
 QY 1201 CTGTCCGTGTATCTCCACAGCAGGGCCCGCGAGCTCCCGCGCTGCCCGCTACCAC 1260
 DB 377 -----ProProThrSer-----GlyMetProProThr----- 386
 QY 1261 CCCTTCACTACCCACTCCGATACTTCTCCAGCTGTACCCCATCACAGTTTGGCGCG 1320
 DB 387 -----SerGlyProProSerAlaMetThrPro-----Pro 396
 QY 1321 TTCCCGCATCATCACTTATCCAGACGAGAGATTGTGAATCTCTTTCATCCCAACCCAG 1380
 DB 397 TyrProGlnPheGluGln---SerGluThrGluThrValHisLeuPheIleProAlaLeu 415
 QY 1381 GCTGTGGCGCCCATCATCGGGAAGAGGGGCGACACATCAAAACAGCTGGCGAGATTCCGC 1440
 DB 416 SerValGlyAlaIleIleGlyLysGlnGlyGlnHisIleLysGlnLeuSerArgPheAla 435
 QY 1441 GGAGCTCTATCAAGATTGCCCTCGGAGGCCAGACCTCAGCGAAAGAGTGTTCATC 1500
 DB 436 GlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLysValArgMetValIle 455
 QY 1501 ATCACCGGGCCACCGGAAGCCAGTTCAGAGCCCGAGGAGCGATCTTTGGGAAACTGAA 1560
 DB 456 IleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrrGlyLysIleLys 475
 QY 1561 GAGGAAACCTCTTTTAAACCCCAAGAGAGTGAAGTGAAGCGCATATCAAGTGC 1620
 DB 476 GluGluAsnPheValSerProLysGluValLysLeuGluAlaHisIleArgValPro 495
 QY 1621 TCTTCCACAGCTGGCGGGTGTATGGCAAGAGTGGAGCGGAGCCGCTGACAGAACTGCAGAAC 1680
 DB 496 SerPheAlaAlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsn 515
 QY 1681 TTAACAGATGCAGAAATCATCTGCTGTCGACCAAGCGCAGATGAAATATGAGGAGTG 1740
 DB 516 LeuSerSerAlaGluValValProArgAspGlnThrProAspGluAsnAspGlnVal 535

1741 ATCGTCAGAAATTATCGGCACCTCTTGTCTAGCCAGACTGCACAGCGCAAGATCAGGAA 1800
 ::
 536 ValValLysIleThrGlyHisPheYrAlaCysGlnValAlaGlnArgLysIleGlnGlu 555
 ::
 1801 ATTGTACAAACAGGTGAAGCAGCAGGAGCAGAAA 1833
 ::
 556 IleLeuThrGlnValLysGlnHisGlnGlnGln 566
 ::
 RESULT 14
 ASP61973
 ID ABP61973 standard; protein; 579 AA.
 AC ABP61973;
 XX
 XX
 DT 07-OCT-2002 (first entry)
 XX
 XX Human lung cancer associated protein sequence SEQ ID NO:446.
 DE Human; lung cancer; lung tumour; cytostatic; gene therapy; vaccine.
 KW Human; lung cancer; lung tumour; cytostatic; gene therapy; vaccine.
 XX Homo sapiens.
 OS
 XX WO200247534-A2.
 PN
 XX 20-JUN-2002.
 PD
 XX 30-NOV-2001; 2001WO-US047576.
 PF
 XX 12-DEC-2000; 2000US-00735705.
 PR 07-MAY-2001; 2001US-00850716.
 PR 28-JUN-2001; 2001US-00897778.
 XX
 XX (CORI-) CORIXA CORP.
 PA
 XX Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
 PI McNeill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedvick TS;
 PI Carter D, Watanabe Y, Peckham DW, Cai F, Foy TW;
 XX
 XX WPI: 2002-583465/62.
 DR N-PSDB; ABQ92483.
 XX
 XX Novel lung carcinoma polynucleotide sequences and polypeptides encoded by
 PT the polynucleotides, useful in pharmaceutical compositions such as
 PT vaccines and as markers to indicate the presence of lung cancer.
 XX
 PS Claim 9; Page 372-374; 381pp; English.
 XX
 CC The present invention describes isolated human lung carcinoma
 CC polynucleotides (I) and polypeptides (II). (I) and (II) have cytostatic
 CC activity, and can be used in gene therapy and in vaccines. Compositions
 CC comprising (I) or (II) can be used for stimulating an immune response in
 CC a patient and for treating lung cancer in a patient. Oligonucleotides of
 CC (I) can be used for detecting the presence of a cancer in a patient, by
 CC obtaining a biological sample from the patient, contacting the biological
 CC sample with the oligonucleotide, detecting in the sample, an amount of
 CC polynucleotide that hybridises to the oligonucleotide and comparing the
 CC amount of polynucleotide that hybridises to the oligonucleotide to a
 CC predetermined cut-off value, and determining the presence of a cancer in
 CC the patient. (I) and (II) are useful in pharmaceutical compositions, e.g.
 CC vaccines. (II) is useful as a marker to indicate the presence or absence
 CC of a cancer such as lung cancer. ABQ92145 to ABQ92486 and ABP61866 to
 CC ABP61992 represent sequences used in the exemplification of the present
 CC invention
 XX
 SQ Sequence 579 AA;
 XX
 Alignment Scores:
 Pred. No.: 6,18e-168 Length: 579
 Score: 1957.50 Matches: 387
 Percent Similarity: 79.19% Conservative: 81
 Best Local Similarity: 65.48% Mismatches: 94
 Query Match: 31.98% Indels: 29

DB: 5 Gaps: 8
 US-09-270-437D-6 (1-3412) x ABP61973 (1-579)
 QY 73 ATGAACAAGCTTTACATCGGAACCTGAGCCCGCCGCTCACCGCGACGACCTCCGGCAG 132
 Db 1 MetAsnLysLeuTyrIleGlyAsnLeuSerGluAsnAlaAlaProSerAspLeuGluSer 20
 QY 133 CTCCTTTGGGACAGGAAGCTGCCCTCGGGGACAGGTCCTGCTGAAGTCGCGGTACGCG 192
 Db 21 IlePheLysAspAlaIleProValSerGlyProPheLeuValLysThrGlyTyrAla 40
 QY 193 TTCGTGGACTACCCCGACAGAACTGGGCCATCCCGCGCATCGACACCTCTCGGGTAAA 252
 Db 41 PheValAspCysProAspGluSerIrpAlaLeuLysAlaIleGluAlaLeuSerGlyLys 60
 QY 253 GTGGAATTGCATGGGAAAATCATGGAAGTTGATTACTCAGTCTCTTAAAAAGTAAAGAGC 312
 Db 61 IleGluLeuHisGlyLysProIleGluValGluHisSerValProLysArgGlnArgIle 80
 QY 313 AGGAAATTCAGATTCGAAACATCCTCCTCACCTGCGAGTGGAGGTCTTGGATGAGCTT 372
 Db 81 ArgLysLeuGlnIleArgAsnIleProProHisLeuGlnIrpGluValLeuAspSerLeu 100
 QY 373 TTGGCTCAATATGGGACAGTGGAGAACTGTGGAACAAGTCAACACACACACAGAAACCGCC 432
 Db 101 LeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAspSerGluThrAla 120
 QY 433 GTTGTCAACGTCAATATGCAACAGAGAGAAGCAAAATAGCCATGGAGAGTAAAGC 492
 Db 121 ValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn 140
 QY 493 GGGCATCAGTTTGGAGAACTACTCCTTCAAGATTTCTACATCCCGGATCAAGAGGTGAGC 552
 Db 141 GlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAspGluThrAlaAla 160
 QY 553 TCCCTTCGCCCCCTCAG-----CGAGCCAGCGT-----GGGAGCCACTCTTCCCGG 600
 Db 161 GlnGlnAsnProLeuGlnGlnProArgGlyArgGlyLeuGlyGlnArgGlySerSer 180
 QY 601 GAGCAAGGCCACGCCCTGGGGGCACCTTCTCAGGCCACAGACAGATTGTCCTCGGTGGG 660
 Db 181 ArgGlnGly--SerProGlySerValSerLysGlnLysProCysAspLeuProLeuArg 199
 QY 661 ATCTGTGTCGCCACCCAGTCTTGTGTCCTCATCGGAAAGGAGGGCTTCACCATTAAG 720
 Db 200 LeuLeuValProThrGlnPheValGlyAlaIleIleGlyLysGlnGlyAlaThrIleArg 219
 QY 721 AACATCAATAGCAGACCCAGTCCCGGTAGATATCCATAGAAAAGAGAACTCTGGAGCT 780
 Db 220 AsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGlyAla 239
 QY 781 GCAGAGAGCTGTCAACATCCATGCCACCCAGAGGGGACTTCTGAAGCATGCGCGCATG 840
 Db 240 AlaGluLysSerIleThrIleLeuSerThrProGluGlyThrSerAlaAlaCysLysSer 259
 QY 841 ATTCTTGAATCATCCAGAAAGAGGCACATGAGACCAACTAGCCGGAAGAGATTCCTCTG 900
 Db 260 IleLeuGluIleMetHisLysGluAlaGlnAspIleLysPheThrGluGluIleProLeu 279
 QY 901 AAAATCTTTGGCACAAATGGCTTGGTTGGAAGACTGATTGGAAGAAGAGGAGCAAAATTG 960
 Db 280 LysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsnLeu 299
 QY 961 AAGAAATGTACATGAACAGGAGGACCAAGATAACAATCTCATCTTTCAGAGATTGAGC 1020
 Db 300 LysLysIleGluGlnAspThrAspThrLysIleThrIleSerProLeuGlnGluLeuThr 319
 QY 1021 ATATACAACCCGGAAGAAACCATCACTGTGAAGGCGCACAGTTGAGCGCTGTCCAGTGT 1080
 Db 320 LeuTyrAsnProGluArgThrIleThrValLysGlyAsnValGluThrCysAlaLysAla 339
 QY 1081 GAGATAGAGATTATGAGAGAGCTGCGTGGAGGCTTTGAAATATGATATATGCTGCTGTTAAC 1140


```

Db      340  GluGluGluLeuMetLysLysIleArgGluSerTyrGluAsnAspIleAlaSerMetAsn 359
Qy      1141 CAACAAGCAATCTGATCCAGGGTTGAACCTCAGCGCACTTGGCATCTTTTCAACAGGA 1200
Db      360  LeuGlnAlaHisLeuIleProGlyLeuAsnLeuAlaLeuGlyLeuPhe----- 376
Qy      1201  CTGTCCTGTGCTATCTCCACAGCAGGCGCCCGGAGCTCCCGGCTGCCCGCTACCCCTACCAC 1260
Db      377  -----ProProThrSer-----GlyMetProProProThr----- 386
Qy      1261  CCCTTCACTCCCACTCCGGATCTTCTCCAGCGCTGTACCCCACTCACCAGTTTGGGCGG 1320
Db      387  -----SerGlyProProSerAlaMetThrPro-----Pro 396
Qy      1321  TTCCCGCATCATCTCTTATCCAGACAGAGATTGTGAATCTCTTCATCCCAACCCAG 1380
Db      397  TyrProGlnPheGluGln--SerGluThrGluThrValHisLeuPheIleProAlaLeu 415
Qy      1381  GCTGTGGCGGCATCATCGGGAAGAGGGGCACACATCAACAGCTGGCGAGTTCCGCC 1440
Db      416  SerValGlyAlaIleIleGlyLysGlnGlnHisIleLysGlnLeuSerArgPheAla 435
Qy      1441  GGAGCCTCTATCAAGATTGCCCTGCGGAAGGCGCCAGACGTCAGCGAAGGATGCTCATC 1500
Db      436  GlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLysValArgMetValIle 455
Qy      1501  ATCAGCGGGCACCGGAAGCCAGATTCAAGGCCAGGAGCGGATCTTTGGGAACCTGAAA 1560
Db      456  IleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysIleLys 475
Qy      1561  GAGGAACACTCTTTAAACCCCAAGAGAGTGAAGCTGGAAGCGCATATCAGAGTCCGC 1620
Db      476  GluGluAsnPheValSerProLysGluGluValLysLeuGluAlaHisIleArgValPro 495
Qy      1621  TCTTCCACAGTGGCGCGGTGATTGGCAAGAGTGGCAAGCCGTGAACGAACTGCAGAAC 1680
Db      496  SerPheAlaAlaGlyArgValIleGlyLysGlyGlyLysThrValAsnGluLeuGlnAsn 515
Qy      1681  TTAACAGTGCAGAAATCATCGTCCGTGACCAACCGCAGATGAAATAGAGAAAGTG 1740
Db      516  LeuSerSerAlaGluValValProArgAspGlnThrProAspGluAsnAspGlnVal 535
Qy      1741  ATCGTCAAGATTATCGGCGACTCTTCTGCTAGCCAGACTGCACAGCGCAGATCAGGAA 1800
Db      536  ValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGlnArgLysIleGlnGlu 555
Qy      1801  ATTGTACAACAGGTGAAGCAGCAGGACAGAAA 1833
Db      556  IleLeuThrGlnValLysGlnHisGlnGlnGln 566
RESULT 15
ID      ADA28536
XX      AC      ADA28536;
XX      DT      20-NOV-2003 (first entry)
XX      DE      Recombinant human lung tumour protein I523S #1.
XX      KW      cancer; lung cancer; gene therapy; vaccine; human;
XX      KW      lung squamous cell carcinoma.
XX      OS      Homo sapiens.
XX      FN      US2003064947-A1.
XX      PD      03-APR-2003.
XX      PF      30-NOV-2001; 2001US-00007700.
XX      PR      18-MAR-1998; 98US-00040802.

```

```

PR      27-JUL-1998; 98US-00123912.
PR      22-DEC-1998; 98US-00221107.
PR      02-APR-1999; 99US-00285479.
PR      17-DEC-1999; 99US-00466396.
PR      30-DEC-1999; 99US-00476496.
PR      10-JAN-2000; 2000US-00480884.
PR      22-FEB-2000; 2000US-00510376.
PR      04-APR-2000; 2000US-00542615.
PR      28-JUN-2000; 2000US-00606421.
PR      02-AUG-2000; 2000US-00630940.
PR      21-AUG-2000; 2000US-00643597.
PR      13-SEP-2000; 2000US-00662786.
PR      09-OCT-2000; 2000US-00685696.
PR      12-DEC-2000; 2000US-00735705.
PR      07-MAY-2001; 2001US-00850716.
PR      28-JUN-2001; 2001US-00897778.
XX
XX      (CORI-) CORIXA CORP.
XX
XX      Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA,
XX      McNeill PD, Fanger N, Retter MW, Durham M, Fanger GK, Vedwick TS,
XX      Carter D, Watanabe Y, Peckham DW, Cai F, Foy TM;
XX      WPI; 2003-540798/51.
XX
XX      New isolated polynucleotides and polypeptides useful for diagnosing,
XX      preventing and/or treating cancer, particularly lung cancer.
XX      Claim 9; Page 285-287; 296pp; English.
XX
XX      The invention describes isolated polynucleotides and polypeptides useful
XX      for diagnosing, preventing and/or treating cancer, particularly lung
XX      cancer. A new isolated polynucleotide comprises: any of the 22 fully
XX      defined nucleotide sequences (e.g. 1012, 900 or 2773 bp) given in the
XX      specification; complements of the nucleotide sequences cited above; at
XX      least 10 contiguous residues of the nucleotide sequences cited above; a
XX      sequence that hybridise to any of the nucleotide sequences under highly
XX      stringent conditions; a sequence that is at least 75 or 90% identical to
XX      the above nucleotide sequences; or degenerate variants of the above
XX      nucleotide sequences. The composition and methods are useful in
XX      diagnosing, preventing and/or treating cancer, particularly lung cancer,
XX      in gene therapy and in vaccines. This is the amino acid sequence of a
XX      recombinant human lung tumour associated protein.
XX      Sequence 579 AA;
SQ

```

Alignment Scores:
Pred. No.: 6.18e-168 Length: 579
Score: 1957.50 Matches: 387
Percent Similarity: 79.19% Conservative: 81
Best Local Similarity: 65.48% Mismatches: 94
Query Match: 31.98% Indels: 29
DB: 7 Gaps: 8

US-09-270-437D-6 (1-3412) x ADA28536 (1-579)

```

Qy      73  ATGAACAAGCTTTACATCGGGAACCTGAGCCCGCCCTCACCAGCGAGCACTTCGGCAG 132
Db      1  MetAsnLysLeuTyrlleGlyAsnLeuSerGluAsnAlaAlaProSerAspLeuGluSer 20
Qy      133  CTCCTTTGGGACAGGAGCTGCCCTGGCGGGACAGGTCTCTGTAAGTCCGGCTACGCC 192
Db      21  IlePheLysAspAlaLysIleProValSerGlyProPheLeuValLysThrGlyTyrAla 40
Qy      193  TTCGTGGACTACCCCGACCACTGGGCCATCGCGCCATCGAGACCCCTCTCGGGTAAA 252
Db      41  PheValAspCysProAspGluSerTrpAlaLeuLysAlaIleGluAlaLeuSerGlyLys 60
Qy      253  GTGAATTCATGGGAAAATCATGGAAGTTGATTCTACTGCTCTCTAAAAGCTAAGAGC 312
Db      61  IleGluLeuHisGlyLysProIleGluValGluHisSerValProLysArgGlnArgIle 80
Qy      313  AGGAATTCAGATTTCGAACATCCCTCCTCACCCTGCGGAGGTGTGGATGGACTT 372

```

```

Db      81  ArgLysLeuGlnIleArgAsnIleProHisLeuGlnTipGluValLeuAspSerLeu 100
QY      373  TTGGCTCAATATGGGACAGTGGAGATGGACAGTCAACACACAGACACAGAAACCGCC 432
Db      101  LeuValGlnTyrGlyValGluSerCysGluGlnValAsnThrAspSerGluThrAla 120
QY      433  GTTGTCAACGTCACATATCAACACAGAGAAAGCAAAATAGCCATGAGAGCTTAAGC 492
Db      121  ValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn 140
QY      493  GGGCATCAGTTTGAAGACTTCCTCAAGATTCTCATCCCGATGAAGAGGTGAGC 552
Db      141  GlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAspGluThrAlaAla 160
QY      553  TCCCTTCGCGCCCTCAG-----CGAGCCAGCGT-----GGGACCACCTCTCCCGG 600
Db      161  GlnGlnAsnProLeuGlnProArgGlyArgGlyLeuGlyGlnArgGlySerSer 180
QY      601  GAGCAAGGCACGCCCCCTGGGGGCACCTCTCAGGCCACAGACAGATTGATTCCTCGG 660
Db      181  ArgGlnGly--SerProGlySerValSerLysGlnLysProCysAspLeuProLeuArg 199
QY      661  ATCCTGGTCCCCACCCAGTTTGTGGTCCCATCATCGGAAAGAGGGCTTGACCATAAAG 720
Db      200  LeuLeuValProThrGlnPheValGlyAlaIleGlyLysGluGlyAlaThrIleArg 219
QY      721  AACATCACTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAGAGAACTCTCGAGCT 780
Db      220  AsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGlyAla 239
QY      781  GCAGAAAGCCTGTCCACATCCATGCCACCCAGAGGGGACTTCTGAAGCATCGCCGATG 840
Db      240  AlaGluLysSerIleThrIleLeuSerThrProGluGlyThrSerAlaAlaCysLysSer 259
QY      841  ATTCTTGAATCATCGAAGAGGACAGATGAGACCAACTAGCCAGAGATTCTCTCTG 900
Db      260  IleLeuGluIleMetHisLysGluAlaGlnAspIleLysPheThrGluGluIleProLeu 279
QY      901  AAAATCTTGGCACAAATGGCTTGTGGAGACTGATTGGAAAGAGAGGACAGAAATTG 960
Db      280  LysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsnLeu 299
QY      961  AAGAAATTTGAACATGAACAGGGACCAAGATACAAATCTCATCTTTGAGGATTGAGC 1020
Db      300  LysLysIleGluGlnAspThrAspThrLysIleThrIleSerProLeuGlnGluLeuThr 319
QY      1021  ATATACACCCGGAAAGACCATCACTGTGAAGGCACAGTTGAGGCCTGTGCCAGTGT 1080
Db      320  LeuTyrAsnProGluArgThrIleThrValLysGlyAsnValGluThrCysAlaLysAla 339
QY      1081  GAGATAGAGATTATGAAGAAGCTGCGTAGGGCTTTGAAATGATATGCTGCTGTTAAC 1140
Db      340  GluGluGluIleMetLysLysIleArgGluSerTyrGluAsnAspIleAlaSerMetAsn 359
QY      1141  CAACAAAGCAATCTGATCCCGGGTTGAACCTCAGCGCATTTGGCATCTTTTCAACAGGA 1200
Db      360  LeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeuGlyLeuPhe----- 376
QY      1201  CTGTCCGTGCTATCTCCACACAGAGGGCCCGCGAGCTCCCGCGCTGCCCGCTTACCAC 1260
Db      377  -----ProProThrSer-----GlyMetProProThr----- 386
QY      1261  CCTTCACTACCCATCCGGATACCTTCTCCAGCCTGTACCCCATCAGGTTTGGCCCG 1320
Db      387  -----SerGlyProProSerAlaMetThrPro-----Pro 396
QY      1321  TTCCCGCATCATCACTCTTATCCAGACGAGAGATTGTGAATCTCTTATCCCAACCCAG 1380
Db      397  TyrProGlnPheGluGln---SerGluThrValHisLeuPheIleProAlaLeu 415
QY      1381  GCTGTGGCGGCATCATCGGGAAGAGGGGGGCACATCAACAGCTGGCGAGATTGCGCC 1440

```

```

Db      416  SerValGlyAlaIleIleGlyLysGlnGlyGlnHisIleLysGlnLeuSerArgPheAla 435
QY      1441  CGAGCCTCTATCAAGATTGCCCTCGGGAAGCCCGAGAGCTCAGCGAAAGGATGTCATC 1500
Db      436  GlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLysValArgMetValIle 455
QY      1501  ATCACCAGGCGCCACCCGGAAGCCAGTTCAAGGCCAGGAGCGGATCTTTGGGAAACTGAAA 1560
Db      456  IleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysIleLys 475
QY      1561  GAGGAAAACCTCTTTTAAACCCCAAGAAAGTGAAGTGGAGCCCATATCATAGAGTGCCC 1620
Db      476  GluGluAsnPheValSerProLysGluGluValLysLeuGluAlaHisIleArgValPro 495
QY      1621  TCTTCCACAGCTGGCGGGTGTGTCGCAAGGTGGCAAGCCGTGAACGAACTGCAGAAC 1680
Db      496  SerPheAlaAlaGlyArgValIleGlyLysGlyGlyLysThrValAsnGluLeuGlnAsn 515
QY      1681  TTAACCAAGTCAGAAAGTTCATCGTCTCGTACCAAAAGCCAGATGAAATGAGGAAGTG 1740
Db      516  LeuSerSerAlaGluValValProArgAspGlnThrProAspGluAsnAspGlnVal 535
QY      1741  ATCGTCAGAAATATCGGGCACTTCTTGTAGCCAGACTGCACAGCGCAGATCAGGGAA 1800
Db      536  ValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGlnArgLysIleGlnGlu 555
QY      1801  ATTGTACAAACAGGTGAAGCAGCAGGAGCAGAAA 1833
Db      556  IleLeuThrGlnValLysGlnHisGlnGlnGln 566

```

Search completed: July 16, 2004, 10:38:50
Job time : 191 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: July 16, 2004, 10:30:59 ; Search time 33 Seconds
(without alignments)

10675.625 Million cell updates/sec

Title: US-09-270-437D-6

Perfect score: 6121
Sequence: 1 ggcagcgaggagcgaggga.....aacctgaaatgtttattt 3412

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 778828

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US09270437/runat 16072004 113050 13370/app query.fasta_1.3591
-DB=Issued Patents AA -OPMT=fastan -SUFFIX=rai -MINMATCH=0 1 -LOOPC1=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09270437 @CGN 1.1 46 @runat 16072004 113050 13370 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PTCUTS.COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2019.5	33.0	577	3	US-09-261-855-2
2	1957.5	32.0	579	4	US-09-643-597-348
3	1957.5	32.0	579	4	US-09-542-615A-348
4	1957.5	32.0	579	4	US-09-606-421B-348
5	1950.5	31.9	579	4	US-09-643-597-176
6	1950.5	31.9	579	4	US-09-480-884A-176
7	1950.5	31.9	579	4	US-09-542-615A-176
8	1950.5	31.9	579	4	US-09-606-421B-176
9	245	4.0	644	1	US-08-021-608D-2
10	245	4.0	644	1	US-08-726-160-2
11	245	4.0	644	5	PCT-US94-01782-2
12	243.5	4.0	643	1	US-08-021-608D-10

13	243.5	4.0	643	1	US-08-726-160-10	Sequence 10, Appl
14	243.5	4.0	643	5	PCT-US94-01782-10	Sequence 10, Appl
15	236.5	3.9	590	1	US-08-021-608D-8	Sequence 8, Appl
16	236.5	3.9	590	1	US-08-726-160-8	Sequence 8, Appl
17	236.5	3.9	590	5	PCT-US94-01782-8	Sequence 8, Appl
18	226	3.7	530	1	US-08-187-793-4	Sequence 4, Appl
19	215	3.5	48	3	US-09-261-855-20	Sequence 20, Appl
20	210	3.4	48	3	US-09-261-855-24	Sequence 24, Appl
21	210	3.4	49	3	US-09-261-855-18	Sequence 18, Appl
22	205	3.3	720	4	US-09-252-991A-21881	Sequence 21881, A
23	201	3.3	47	3	US-09-261-855-21	Sequence 21, Appl
24	200	3.3	49	3	US-09-261-855-22	Sequence 22, Appl
25	197	3.2	47	3	US-09-261-855-17	Sequence 17, Appl
26	195	3.2	1706	4	US-09-252-991A-31760	Sequence 31760, A
27	190	3.1	1020	4	US-09-252-991A-28970	Sequence 28870, A
28	186.5	3.0	1476	4	US-09-252-991A-29427	Sequence 29427, A
29	184.5	3.0	1228	4	US-09-252-991A-17764	Sequence 17764, A
30	184	3.0	343	1	US-08-187-793-2	Sequence 2, Appl
31	181	3.0	748	4	US-09-252-991A-18427	Sequence 18427, A
32	177	2.9	47	3	US-09-261-855-23	Sequence 23, Appl
33	176.5	2.9	471	3	US-08-866-928B-1	Sequence 1, Appl
34	176.5	2.9	471	4	US-09-685-836-1	Sequence 1, Appl
35	176	2.9	47	3	US-09-261-855-19	Sequence 19, Appl
36	176	2.9	705	4	US-09-252-991A-30792	Sequence 30792, A
37	175.5	2.9	1002	4	US-09-252-991A-27980	Sequence 27980, A
38	174.5	2.9	1780	1	US-08-769-309A-5	Sequence 5, Appl
39	174.5	2.9	1780	3	US-08-994-570-5	Sequence 5, Appl
40	174	2.8	657	4	US-09-252-991A-28001	Sequence 28001, A
41	173.5	2.8	705	4	US-09-252-991A-23298	Sequence 23298, A
42	172	2.8	719	4	US-09-252-991A-23660	Sequence 23660, A
43	171.5	2.8	697	4	US-09-252-991A-24009	Sequence 24009, A
44	169.5	2.8	2294	4	US-09-252-991A-17231	Sequence 17231, A
45	169	2.8	742	4	US-09-252-991A-32659	Sequence 32659, A

ALIGNMENTS

RESULT 1
US-09-261-855-2
; Sequence 2, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; FILE REFERENCE: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 577
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-261-855-2

Alignment Scores:				
Pred. No.:	1.24e-170	Length:	577	
Score:	2019.50	Matches:	394	
Percent Similarity:	78.94%	Conservative:	82	
Best Local Similarity:	65.34%	Mismatches:	96	
Query Match:	32.99%	Indels:	31	
DB:	3	Gaps:	4	
US-09-270-437D-6 (1-3412) x US-09-261-855-2 (1-577)				

Qy	73	ATGACAGCTTACATCGGAGACCTAGCCGCCCTACCGCCGACGACCTCCGCGAG	132
Db	1	MetAsnLysLeuTyrIleGlyAsnLysLeuSerValThrProAlaAspLeuGluLys	20
Qy	133	CTCTTGGGACAGGAGCTGCCCTCGCGGAGAGCTCTGCTGAAGTCCGCTAGCC	192
Db	21	ValPheAlaGluHisIleSerTyrSerGlyGlnPheLeuValLysSerGlyTyrAla	40

```
QY 193 TTCGTGGACTACCCGACAGAACTGGGCGCATCCGCGCATCGAGACCTCTCGGATAA 252
Db 41 PheValAspCysProAspGluHisThrPalaMetLysAlaIleGluThrPheSerGlyLys 60
QY 253 GTGGAATTGCATGGGAAATCATGGAAGTTGATTACTCAGTCTCTAAAAGCTAAGGAGC 312
Db 61 ValGluLeuGlnGlyArgLeuGluMetGluHisSerValProLysLysGlnArgSer 80
QY 313 AGGAAATTCAGATTCGAAACATCCCTCTCTACCTGCGAGTGGAGGTGTGGATGCATT 372
Db 81 ArgLysIleGlnIleArgAsnIleProProGlnLeuArgTrpGluValLeuAspSerLeu 100
QY 373 TTGGCTCAATATGGGACACTGGAGATGTGGAACAAGTCAACACAGACACAGAAACCGCC 432
Db 101 LeuAlaGlnTyrGlyThrValGluAsnCysGluGlnValAsnThrGluSerGluThrAla 120
QY 433 GTTGTCAACGTCACATATGCAACAAGAGAAAGCAAAATAGCCATGGAGAAGCTAAGC 492
Db 121 ValValAsnValThrTyrSerAsnArgGluGlnThrArgGlnAlaIleMetLysLeuAsn 140
QY 493 GGGCATCATGTTGAGAACTACTCTTCAAGATTTCTACATCCCGCATCAAGAGGTGAGC 552
Db 141 GlyHisGlnLeuGluAsnHisAlaLeuLysValSerTyrIleProAspGluGlnIleThr 160
QY 553 TCCCTTCCGCCCTCAGCGAGCCAGCGTGGGGACCACTCTTCCCGGGAGCAAGGCCAC 612
Db 161 -----GlnGlyProGluAsnGlyArgArgGlyGlyPheGlySerArgGlyGlnProArg 178
QY 613 -----GCCCTCGGGCACTTCTCAGGCCGACAGATTCATTCCTCGCGT 657
Db 179 GlnGlySerProValAlaAlaGlyAlaProAlaLysGlnGlnProValAspIleProLeu 198
QY 658 CGATCTCGTCCCGACCCAGTTTGTGTGTCATCATCGCAAGAGGGCTTGACCAT 717
Db 199 ArgLeuLeuValProThrGlnTyrValGlyAlaIleIleGlyLysGluGlyAlaThrIle 218
QY 718 AAGACATCATCAAGCAGACCCAGTCCCGGTPAGATATCCATAGAAAGAGAACTCTGGA 777
Db 219 ArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGly 238
QY 778 GTGCAGAGAACCTGTGCATCCATCCATGCCACCCAGAGGGGACTTCTGAAGCATCGCCG 837
Db 239 AlaAlaGluLysAlaIleSerValHisSerThrProGluGlyCysSerSerAlaCysLys 258
QY 838 ATGATCTTGAATCATCAGAAAGAGGAGCAGATGAGACCAACTAGCCGCAAGATTCCT 897
Db 259 MetIleLeuGluIleMetHisLysGluAlaLysAspThrLysThrAlaAspGluValPro 278
QY 898 CTGAAATCTTGGCACACAATGGCTTGGTTGGAAGACTGATTGGAAGAAAGAGCAAAAT 957
Db 279 LeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsn 298
QY 958 TTGAAGAAATTTGAATCAACACAGGACCAAGATAACAATCTCATCTTTGAGGATTTG 1017
Db 299 LeuLysLysValGluAspThrGluThrLysIleThrIleSerSerLeuGlnAspLeu 318
QY 1018 AGCATATACACCCGGAAGACCATCACTGTGAAGGAGCAGATGAGGCTGTGCCAGT 1077
Db 319 ThrLeuTyrAsnProGluArgThrIleThrValLysGlyAlaIleGluAsnCysCysArg 338
QY 1078 GCTGAGATGAGATTATGAAGAAGCTGGTGGAGGCTTTGAAATGATATGCTGCTGTT 1137
Db 339 AlaGluGlnGluIleMetLysLysValArgGluAlaTyrGluAsnAspValAlaAlaMet 358
QY 1138 AACCAACAGCAATCTCATCCAGGTTTGAACCTCAGGCACTTGGCATCTTTTCAACA 1197
Db 359 SerLeuGlnSerHisIleProGlyLeuAsnLeuAlaAlaValGlyLeuPheProAla 378
QY 1198 GGACTGTCCGTCTATCTCCACAGCAGGGGCCCGGGAGCTCCCGCGCTCCCGCTAC 1257
Db 379 SerSerAlaValProPro-----ProSerSerValThrGlyAlaAlaProTyr 396

RESULT 2
US-09-643-597-348
; Sequence 348, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 348
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-643-597-348
```

Alignment Scores:

Pred. No.: 4, 02e-165 Length: 579
 Score: 1957.50 Matches: 367
 Percent Similarity: 79.19% Conservative: 81
 Best Local Similarity: 65.48% Mismatches: 94
 Query Match: 31.98% Indels: 29
 DB: 4 Gaps: 8

US-09-270-437D-6 (1-3412) x US-09-643-597-348 (1-579)

```

Qy 73 ATGAACAGCTTACATCGGAACCTGAGCCCGCCGTCACCGCGACGACCTCCGGCAG 132
Db 1 MetAsnLysLeuTyrlleGlyAsnLeuSerGluAsnAlaAlaProSerAspLeuGluSer 20

Qy 133 CTCCTTGGGACAGGAGCTCCCTGCGGACAGGCTCCTGCTGAAGTCCGGTACGCC 192
Db 21 IlePheLysAspAlaLysIleProValSerGlyProPheLeuValLysThrGlyTyra 40

Qy 193 TTCGTGGACTACCCGACAGAACTGGGCCCATCCGCGCCATCGAGACCTCTCGGTAAA 252
Db 41 PheValAspCysProAspGluSerTrpAlaLeuLysAlaIleGluAlaLeuSerGlyLys 60

Qy 253 GTGGAATTCATGGAAATCATGGAAGTGAATCTAGTCTCTAAAGACTAAGAGC 312
Db 61 IleGluLeuHisGlyLysProIleGluValGluHisSerValProLysArgGluArgIle 80

Qy 313 AGGAAATTCAGATTGAAACATCCCTCCCTCACCTGAGTGGGAGGTGTGGATGAGCTT 372
Db 81 ArgLysLeuGlnIleArgAsnIleProPheHisLeuGlnTrpGluValLeuAspSerLeu 100

Qy 373 TTGGCTCAATATGGGACAGTGGAGAATGTGAACAAGTCAACACAGACACAGAAACCGCC 432
Db 101 LeuValGlnTyrglyValValGluSerCysGluGlnValAsnThrAspSerGluThra 120

Qy 433 GTTGTCAAGTCATATGCAACAGAGAGAGAACAAATAGCCATGGAGAACTAAGC 492
Db 121 ValValAsnValThrTyrsSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn 140

Qy 493 GGGCATCAGTTTGAGAACTACTCTCAAGATTTCCTACATCCCGGATGAGAGGTAGC 552
Db 141 GlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrlleProAspGluThraAla 160

Qy 553 TCCCTCTCGCCCTCAG-----CGAGCCCGAGCGT-----GGGACCACTCTTCCCGG 600
Db 161 GlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlyGlnArgGlySerSer 180

Qy 601 GAGCAAGCCACGCCCTCGGGGACACTTCTCAGGCCAGACAGATGTATTTCCCGCTCGG 660
Db 181 ArgGlnGly---SerProGlySerValSerLysGlnLysProCysAspLeuProLeuArg 199

Qy 661 ATCTGTCTCCCCACCCAGTCTGTTGGTGCCATCATCGGAAGGAGGCTTGACCATTAAG 720
Db 200 LeuLeuValProThrGlnPheValGlyAlaIleIleGlyLysGluGlyAlaThrIleArg 219

Qy 721 AACATCACTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAAGAACTCTGGAGCT 780
Db 220 AsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGlyAla 239

Qy 781 GCAGAGAGCCTGTACCATTCATCCACCCAGGAGGACTCTCTGAAGCATCCCGCATG 840
Db 240 AlaGluLysSerIleThrIleLeuSerThrProGluGlyThrSerAlaAlaCysLysSer 259

Qy 841 ATTCTTGAATCATGCAAGAGGAGGATGAGACCAACTAGCCGAGAGATTCCTCTG 900
Db 260 IleLeuGluIleMethIshLysGluAlaGlnAspIleLysPheThrGluGluIleProLeu 279

Qy 901 AAAATCTTGGCACACATGCTTGGTGGTGAAGACTGATTGAAAAGAGGACAGAAATTTG 960
Db 280 LysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgGlnLeu 299

Qy 961 AAGAAATTCAGATGAACAGGAGCAGCAATCAATCTCATCTTTGAGGATTTGAGC 1020
Db 300 LysLysIleGluGlnAspThrAspThrLysIleThrLysIleThrIleSerProLeuGlnLeuThr 319

```

RESULT 3

US-09-542-615A-348

; Sequence 348, Application US/09542615A

; Patent No. 6519256

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; APPLICANT: Fan, Ligu

; APPLICANT: Kalos, Michael D.

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: Hosken, Nancy A.

; APPLICANT: Fanger, Gary R.

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY

; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER

; FILE REFERENCE: 210121.455C8

; CURRENT APPLICATION NUMBER: US/09/542,615A

; CURRENT FILING DATE: 2000-04-14

```

Qy 1021 ATATACAAACCGGAAAGAACCATCACTGTGAAGGGCACAGTTGAGCCCTGTGCCAGTGT 1080
Db 320 LeuTyrsAsnProGluArgThrIleThrValLysGlyAsnValGluThrCysAlaLysAla 339

Qy 1081 GAGATAGAGATATGAAGAAGCTGGTGAGCCCTTGAATATGATATCTGGCTCTTAAC 1140
Db 340 GluGluGluIleMetLysLysIleArgGluSerTyrgluAsnAspIleAlaSerMetAsn 359

1141 CAACAAGCAATCTGATCCCAAGGTGAACCTCAGCGCACITGGCATCTTTTCAACAGA 1200
Db 360 LeuGlnAlaHisLeuIleProGlyLeuAsnLeuAlaLeuGlyLeuPhe----- 376

1201 CTGTCGTGTATCTCCACAGAGGGCCCGCGAGCTCCCGCGTCCCGCCCTTACCAC 1260
Db 377 -----ProThrSer-----GlyMetProProProThr----- 386

1261 CCCTTCACTACCACTCGGNATCTTCTCCAGCTGTACCCCATCACCAGTTTGGCCCG 1320
Db 387 -----SerGlyProProSerAlaMetThrPro-----Pro 396

1321 TTCCCGCATCATCTCTATCCAGAGGAGATTTGAATCTCTTCATCCCAACCCAG 1380
Db 397 TyrProGlnPheGluGln---SerGluThrGluThrValHisLeuPheIleProAlaLeu 415

1381 GCTGTGGCGCCATCATCGGGAAGAGGGGCACACATCAACAGCTGGCGAGATTCGCC 1440
Db 416 SerValGlyAlaIleIleGlyLysGlnGlyGlnHisIleLysGlnLeuSerArgPheAla 435

1441 GGAGCCTCATCAGATTGCCCTGCGGAGGCCACACGTCAGCGAAGAGATGGTCAATC 1500
Db 436 GlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLysValArgMetValIle 455

1501 ATCAACCGGGCACCGGAAGCCAGTTTCAAGCCCGAGGAGCGATCTTTGGGAACTCAA 1560
Db 456 IleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrglyLysIleLys 475

1561 GAGAAACTTCTTTAAACCCCAAGAAAGTGAAGCTGGAAGCGCATATCAGAGTCCCC 1620
Db 476 GluGluAsnPheValSerProLysGluGluValLysLeuGluAlaHisIleArgValPro 495

1621 TCTTCCACAGCTGGCGGTGATTGGCAAGGTGGCAAGACCGTGAACGAACTGCAGAAC 1680
Db 496 SerPheAlaAlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsn 515

1681 TTAACCAAGTGCAGAAGTCACTGTCCTCGTCCGTAACCAACGACAGTGAATAATGAGAA 1740
Db 516 LeuSerSerAlaGluValValProArgAspGlnThrProAspGluAsnAspGlnVal 535

1741 ATCGTCAGAAATTCGGGCACCTCTTTGTAGCCAGACTCCACAGCGCAAGATCAGGAA 1800
Db 536 ValValLysIleThrGlyHisPheTyraLysGlnValAlaGlnArgLysIleGlnGlu 555

1801 ATTGTACAACAGCTGAAGCAGCAGGAGCAGAAA 1833
Db 556 IleLeuThrGlnValLysGlnHisGlnGlnGln 566

```

; NUMBER OF SEQ ID NOS: 350
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 348
 ; LENGTH: 579
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-542-615A-348

Alignment Scores:
 Pred. No.: 4.02e-165 Length: 579
 Score: 1957.50 Matches: 387
 Percent Similarity: 79.19% Conservative: 81
 Best Local Similarity: 65.48% Mismatches: 94
 Query Match: 31.98% Indels: 29
 DB: 4 Gaps: 8

US-09-270-437D-6 (1-3412) x US-09-542-615A-348 (1-579)

QY 73 ATGAACAGCTTTTACATCGGAACCTGAGCCCGCGCTCACCGCCGACGACCTCCGGCAG 132
 Db 1 MetAsnLysLeuTyrlleGlyAsnLeuSerGluAsnAlaA1aProSerAspLeuGluSer 20
 QY 133 CTCCTTTGGGGACAGGAAGTCCCTCGGGGACAGGTCCTGCTGAAGTCGCGGTACGCC 192
 Db 21 IlePheLysAspAlaLysIleProValSerGlyProPheLeuValLysThrGlyTyrAla 40
 QY 193 TTCGTGGACTACCCGACGACCACTGGCCATCCCGGCCATCGAGACCTCTCCGGTAAA 252
 Db 41 PheValAspCysProAspGluSerTrpAlaLeuLysAlaIleGluAlaLeuSerGlyLys 60
 QY 253 GTGGAAATTCATGGAAATCATGGAAGTTGATCTCAGTCTCTAAAAAGTAAGGAGC 312
 Db 61 IleGluLeuHisGlyLysProIleGluValGluHisSerValProLysArgGlnArgIle 80
 QY 313 AGGAAATTCAGATTGCAAACTCCCTCTACCTGCAAGTGGAGGTGTGATGACTT 372
 Db 81 ArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspSerLeu 100
 QY 373 TTGGCTCAATATGGACAGTGGAGATGTGGACAGTCAACACAGACACAGAAACCGCC 432
 Db 101 LeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAspSerGluThrAla 120
 QY 433 GTTGTCAACGTCACATATGCAACAGAGAAAGCAAAATAGCCATGGAGAAGCTAAGC 492
 Db 121 ValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn 140
 QY 493 GGGCATCATGTTGAGAACTATCTCTCAAGATTCCTCATCCGGATGAAGAGTGAGC 552
 Db 141 GlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAspGluThrAlaAla 160
 QY 553 TCCCTCTCCGCCCTCAG-----CGAGCCGACGT-----GGGGACCACTCTCCCGG 600
 Db 161 GlnGlnAsnProLeuGlnGlnProArgGlyArgGlyLeuGlyGlnArgGlySerSer 180
 QY 601 GAGCAGGCGACCGCCCTCGGGGCACTTCTCAGGCCAGACAGATTGATTTCCCGTGGCG 660
 Db 181 ArgGlnGly--SerProGlySerValSerLysGlnLysProCysAspLeuProLeuArg 199
 QY 661 ATCTGTGTCACCCAGCCAGTTTGTGTGTCATCTCGGAAGAGAGGGCTTGACCATAAAG 720
 Db 200 LeuLeuValProThrGlnPheValGlyAlaIleIleGlyLysGluGlyAlaThrIleArg 219
 QY 721 AACATCACTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAAGAGAACTCTGGAGCT 780
 Db 220 AsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGlyAla 239
 QY 781 GCAGAGAAGCTGTTCACCATCCATGCCACCCAGGGGACTTCTGAAGCATGCCCATG 840
 Db 240 AlaGlyLysSerIleThrIleLeuSerThrProGluGlyThrSerAlaAlaCysLysSer 259
 QY 841 ATTCTTGAATCATGCAAGAGAGGAGGATGAGACCAAACTAGCCAGAGATTCTCTCG 900
 Db 260 IleLeuGluIleMethHisLysGluAlaGlnAspIleLysPheThrGluGluIleProLeu 279

QY 901 AAAATCTTGGCACACAATGGCTTGGTTGGAAGACTGATTGGAAGAAAGGAGGAGAAATTTG 960
 Db 280 LysIleLeuAlaHisAsnAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsnLeu 299
 QY 961 AAGAAAAATTGAACATGAACAGGAGCAAGATAACAATCTCATCTTTCAGAGATTGAGC 1020
 Db 300 LysLysIleGluGlnAspThrAspThrLysIleThrIleSerProLeuGlnGluLeuThr 319
 QY 1021 ATATACAAACCCGGAAGAACCATCACTGTGAAGGSCACAGTTGAGGCTGTGCCAGTCT 1080
 Db 320 LeuTyrAsnProGluArgThrIleThrValLysGlyAsnValGluThrCysAlaLysAla 339
 QY 1081 GAGATAGAGATTATGAAGAAGCTCGTGAGCCCTTTGAAAATGATATGCTGCTGTTAAC 1140
 Db 340 GluGluGluIleMetLysLysIleArgGluSerTyrGluAsnAspIleAlaSerMetAsn 359
 QY 1141 CAACAAGCAATCTGATCCAGGGTTGAACCTCAGCGCAGCTTGGCATCTTTTCAACAGA 1200
 Db 360 LeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeuGlyLeuPhe----- 376
 QY 1201 CTGTCCGTGCTATCTCCACACGACGAGGCCCGCGGAGCTCCCGCGCTGCCCTTACCAC 1260
 Db 377 -----ProProThrSer-----GlyMetProProProThr----- 386
 QY 1261 CCTTCTACTACCCACTCCGGATATCTTCTCCAGCCTGTACCCCAATCACCAGTTTGGCCCG 1320
 Db 387 -----SerGlyProProSerAlaMetThrPro-----Pro 396
 QY 1321 TTCGCGCATCATCTCTTATCCAGAGCAGAGATTGGAATCTCTTATCTCCCAACCCAG 1380
 Db 397 TyrProGlnPheGluGln---SerGluThrValHisLeuPheIleProAlaLeu 415
 QY 1381 GCTGTGGCGCCATCATCGGAAGAGGCGGCACACATCAACAGCTGCGGAGATTGCGC 1440
 Db 416 SerValGlyAlaIleIleGlyLysGlnGlnHisIleLysGlnLeuSerArgPheAla 435
 QY 1441 GGAGCCTCTATCAAGATTGCCCTCGGAAGCCCGACGCTCAGCGAAAGGATGGTCATC 1500
 Db 436 GlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLysValArgMetValIle 455
 QY 1501 ATCAGCGGCGCCAGGAGCCCGAGTTCAGAGCCCGAGGAGCGGATCTTTGGGAAGCTGAA 1560
 Db 456 IleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysIleLys 475
 QY 1561 GAGGAAATCTCTTTAACCCCAAGAGAGTGAAGCTGGAAGCGCATATCAGAGTGCCT 1620
 Db 476 GluGluAsnPheValSerProLysGluGluValLysLeuGluAlaHisIleArgValPro 495
 QY 1621 TCTTCCACAGCTGGCGGGTGTATTCGCAAGAGTGGCAAGACCGTGAACGAACTGCAGAAC 1680
 Db 496 SerPheAlaAlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsn 515
 QY 1681 TTAACAGTGCAGAGTATCATCGTGTCTGTGCAAAAGCCAGATCAAAATCAGGAAGTG 1740
 Db 516 LeuSerSerAlaGluValValProArgAspGlnThrProAspGluAsnGluVal 535
 QY 1741 ATCTGTCAAGATTATCGGCACCTCTTTGTAGCCAGACTGCACAGCGCAAGATCAGGGAA 1800
 Db 536 ValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGlnArgLysIleGlnGlu 555
 QY 1801 ATTGTACACAGCTGAAGCAGGAGGAGCAGAAA 1833
 Db 556 IleLeuThrGlnValLysGlnHisGlnGlnGln 566

RESULT 4
 US-09-606-421B-348
 ; Sequence 348, Application US/09606421B
 ; Patent No. 6531315
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Fan, Liqun
 ; APPLICANT: Kalos, Michael D.

APPLICANT: Bangur, Chaitanya S.	QY	721	AAATCACTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAAGAACTCTGGAGCT	780
APPLICANT: Hosken, Nancy	Db	220	AsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGlyAla	239
APPLICANT: Fanger, Gary R.	QY	781	GCAGAGAAGCCTGTCAACCATCCACCCAGAGGGAGTCTTGAAGCATCCGCGATG	840
APPLICANT: Li, Samuel X.	Db	240	AlaGluLysSerIleThrIleLeuSerThrProGluGlyThrSerAlaAlaCysLysSer	259
APPLICANT: Wang, Aijun	QY	841	ATTCTTGAATCATGTCAGAAAGAGGCGAGATGAGACCAAACTAGCCGAGAGATTCTCTG	900
APPLICANT: Skeiky, Yasir A.W.	Db	260	IleLeuGluIleMetHisLysGluAlaGlnAspIleLysPheThrGluGluIleProLeu	279
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY	QY	901	AAATCTTTGGCACACAATGCGTCTGGTTGGAAGACTGATTGGAAGAAAGAGCAGAAATTTG	960
TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER	Db	280	LysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsnLeu	299
FILE REFERENCE: 210121.455C9	QY	961	AAGAAATTTGAACATGAACAGGACCAAGATTAACATCTCATCTTTTGAGGATTTCGAGC	1020
CURRENT APPLICATION NUMBER: US/09/606,421B	Db	300	LysLysIleGluGlnAspThrAspThrLysIleThrIleSerProLeuGlnGluLeuThr	319
CURRENT FILING DATE: 2000-06-28	QY	1021	ATATACAAACCCGAAAGAACCATCATCTGTGAAGGGCACAGTGTAGGCTGTGCCAGTCT	1080
NUMBER OF SEQ ID NOS: 358	Db	320	LeuTyrAsnProGluArgThrIleThrValLysGlyAsnValGluThrCysAlaLysAla	339
SOFTWARE: FastSeq for Windows Version 3.0	QY	1081	GAGATAGAGATTATGAAGAAGCTCGCTGAGGCCCTTTGAAATATGATATGCTGGCTGTAAAC	1140
SEQ ID NO 348	Db	340	GluGluGluIleMetLysLysIleArgGluSerTyrGluAsnAspIleAlaSerMetAsn	359
LENGTH: 579	QY	1141	CAACAAGCCAAATGTATCCCGAGGTGAACTCAGCGCACTTGGCATCTTTTCAACAGGA	1200
TYPE: PR1	Db	360	LeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeuGlyLeuPhe	376
ORGANISM: Homo sapiens	QY	1201	CTGTCCGTGTATCTCCACGAGGCGCCCGCGGAGCTCCCGCGTCCCGCTACCAC	1260
US-09-606-421B-348	Db	377	-----ProThrSer-----GlyMetProProThr-----	386
Alignment Scores:	QY	1261	CCCTTCACTACCCACTCCGAGTACTTCTCCAGCTGTACCCCATCACAGTTTGGCCCG	1320
Pred. No.: 4.02e-165	Db	387	-----SerGlyProProSerAlaMetThrPro-----	396
Score: 1957.50	QY	1321	TTCCCGCATCATCTCTTATCCAGAGCAGGAGATTGTGAATCTTTCATCCCAACCCAG	1380
Percent Similarity: 79.19%	Db	397	TyrProGlnPheGluGln---SerGluThrGluThrValHisLeuPheIleProAlaLeu	415
Best Local Similarity: 65.48%	QY	1381	GCTGTGGCGCCATCATCGGGAAGAGGGGCGCACATCAACAGCTGGCGAGATTCCGC	1440
Query Match: 31.98%	Db	416	SerValGlyAlaIleIleGlyLysGlnGlyHisIleLysGlnLeuSerArgPheAla	435
DB: 4	QY	1441	GGAGCTCTATCAAGATTGCCCTCGGGAAGGCGCCAGACGCTCAGCGAAAGATGGTTCATC	1500
US-09-270-437D-6 (1-3412) x US-09-606-421B-348 (1-579)	Db	436	GlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLysValArgMetValIle	455
73 ATGAACAAGCTTTACATCGGAACCTGAGCCCCCGCTCACCGCCGAGCACTCCGGCAG	QY	1501	ATCACCGGGCCACCGGAAGCCAGTTCAGGCGCCAGGAGCGGATCTTTGGGAAACTGAAA	1560
1 MetAsnLysLeuTyrIleGlyAsnLeuSerGluAsnAlaAlaProSerAspLeuGluSer	Db	456	IleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysIleLys	475
133 CTCTTTGGGACAGGAAGCTCCCTCGCGGGACAGGTCCTGCTGAAGTCCGGCTACGCC	QY	1561	GAGGAAATCTTCTTAAACCCCAAGAGAGTGAAGCTGAAGCCGATATCAGAGTGC	1620
21 IlePheLysAspAlaLysIleProValSerGlyProPheLeuValLysThrGlyTyrAla	Db	476	GluGluAsnPheValSerProLysGluGluValLysLeuGluAlaHisIleArgValPro	495
193 TTCGTGAGTACTCCCGACAGAACTGGGCCCATCGGCCCATCGAGACCCCTCTCGGTA	QY	1621	TCCTCCACAGCTGGCGGGTGTGGCAAGGTGGCAAGCCGTGAACTGCAGAAC	1680
41 PheValAspCysProAspGluSerTrpAlaLeuLysAlaIleGluAlaLeuSerGlyLys	Db	496	SerPheAlaAlaGlyArgValIleGlyLysGlyThrValAsnGluLeuGlnAsn	515
253 GTGGAAATTCAGTTCGGAACATCCCTCCTCACCTGCAGTGGGAGGTGGTGAATGCACTT	QY	1681	TTAAACAGTGCAGAGTTCATCGTCTCGTACCAACCGCAGATCAAAATAGGAAGTG	1740
61 IleGluLeuHisGlyLysProIleValGluHisSerValProLysArgGlnArgIle	Db	516	LeuSerSerAlaGluValValProArgAspGlnThrProAspGluAsnAspGlnVal	535
313 AGGAAATTCAGATTCCGAAACATCCCTCCTCACCTGCAGTGGGAGGTGGTGAATGCACTT	QY	1741	ATCGTCAGAAATATCGGGCACTCTTTGTAGCCAGCTGCACAGCGCAGATCAGGAA	1800
81 ArgLysLeuGlnIleArgAsnIleProPheIleGluGlnTrpGluValLeuAspSerLeu	Db	536	ValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGlnArgLysIleGlnGlu	555
373 TTGGCTCAATTTGGACAGTGGGAATGTGGAACAAGTCAACACAGACAGAAACCGCC	QY	1801	ATTGTACACAGGTGAGCAGCAGGAGCAGAAA	1833
101 LeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAspSerGluThrAla	Db			
433 GTTGTCACGTCACATATGCAACAGAGACAGAACAAATAGCCATGGAGAGCTAAGC	QY			
121 ValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn	Db			
493 GGGCATCAGTTTGAGAACTACTCTCTCAAGATTTCCTACATCCCGGATGAAGAGGTGAGC	QY			
141 GlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAspGluThrAlaAla	Db			
553 TCCCTTCGCCCTCCAG-----CGAGCCACAGCGT-----GGGGACCACTCTCCCGG	QY			
161 GlnGlnAsnProLeuGlnGlnProArgGlyArgGlyLeuGlyGlnArgGlySerSer	Db			
601 GAGCAAGGCCACGCCCTGGGGCACTTCTCAGGCCACAGATTGATTTCCGCTGGCG	QY			
181 ArgGlnGly---SerProGlySerValSerLysGlnLysProCysAspLeuProLeuArg	Db			
661 ATCTGTGTCGCCACCCAGCTTTGTTGGTCATCATCGGAAAGAGGGCTTGACCATAAAG	QY			
200 leuLeuValProThrGlnPheValGlyAlaIleIleGlyLysGluGlyAlaThrIleArg	Db			

[illegible]

RESULT 7

```

US-09-542-615A-176
; Sequence 176, Application US/09542615A
; Patent No. 6518256
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Pan, Liyun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C8
; CURRENT APPLICATION NUMBER: US/09/542,615A
; CURRENT FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 176
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-542-615A-176

```

Alignment Scores:		
Pred. No.:	1.69e-164	Length:
Score:	1950.50	Matches:
Percent Similarity:	79.02%	Conservative:
Best Local Similarity:	75.31%	Mismatches:
Query Match:	31.67%	Indels:
DB:	4	Gaps:
		8
		29
		85
		386
		579

US-09-270-437D-6 (1-3412) x US-09-542-615A-176 (1-579)

Qy	73	ATGAACAAGCTTTTACATCGGGAACCTGAGACCGCGCCGCTCACGCGCGAGACCTCGCGGAG	132
Db	1	MetAsnIysLeuTyriLeGlyAsnLeuSerGluAsnLaAlaProSerAspLeuGluSer	20
Qy	133	CTCTTTGGGGACAGAAAGCTCCCTCGCGGGACAGGCTCTCAAGTCGGCTACGCC	192
Db	21	IlePheIysAspAlaIysIleProValSerGlyProPheLeuValIysThrGlyTyrAla	40
Qy	193	TTCTGTGGACTACCCCGACCAAGACTGGGCCATTCGCGCCATCGAGACCCCTCTCGGGTAAA	252
Db	41	PheValAspCysProAspGluSerTrpAlaLeuIysAlaIleGluAlaLeuSerGlyIys	60
Qy	253	GTGGAAATTCATGGGAAATCATGGAAGTTGATTACTCAGTCTCTAAAAGCTAAGGAC	312
Db	61	IleGluLeuHisGlyIysProIleGluValGluHisSerValProIysArgGlnArgIle	80

	QY	313	AGGAAAATTACATTCGAAACCATCCTCTCTCACTCGAGTGGCGAGGTGTGCATGACTTT	372
	Dd	81	ArgLysLeuGlnIleArgAsnIleProHisLeuGlnTrpValLeuAspSerLeu	100
	QY	373	TTCGGCTCAATATGGCACAGTAGAGAATGTGGAACAAGTCAACACAGACACAGAACCCGCC	432
	Dd	101	LeuValGlnTyr-GlyValValGluSerCysGluGlnValAsnThrAspSerGluThrAla	120
	QY	433	GTTGTCAAAGCTCACATATGCAACAGAGAAAGCAAAAATAGCCATGGAGAGCTTAAGC	492
	Dd	121	ValValAsnValThr-Tyr-SerSerLysaspGlnAlaArgglnAlaLeuAspLysLeuAsn	140
	QY	493	GGGCATCACTTTGAGAACTACTCTTCAGATTTCTCATCCCCGGATGAAGAGGTGAGC	552
	Dd	141	GlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrlleProAspGluMetAlaAla	160
	QY	553	TCCCCTTCGCCCTTCAG-----CGAGCCACAGCT-----GGGACCACTTCTCCCGG	600
	Dd	161	GlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlyGlnArgGlySerSer	180
	QY	601	GAGCAGGCCACGCCCTCGGGGCACCTCTCAGGCCACAGACATGATTTCCCGTGCCG	660
	Dd	181	ArgGlnGly--SerProGlySerValSerLysGlnLysProCysAspLeuProLeuArg	199
	QY	661	ATCCTCGTCCCAACCAGCTTTGTTGTCGCCATCATCGGAAAGAGGGCTTGACCTAAG	720
	Dd	200	LeuLeuValProThrGlnPheValGlyAlaIleGlyLysGluGlyAlaThrIleArg	219
	QY	721	AACATCACTAAGCAGACCCAGTCGGGTAGATCCATAGAAAGAGAACTCTGGAGCT	780
	Dd	220	AsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGlyAla	239
	QY	781	GCAGAGAAGCCTGTACCATCCATCCACCCAGAGGGGACTTCTGAAGCATGCCCATG	840
	Dd	240	AlaGluLysSerIleThrIleLeuSerThrProGluGlyThr-SerAlaAlaCysLysSer	259
	QY	841	ATTCTTTGAATCATGCAGAAAGAGGCAGATGAGACCAACTAGCCGAGAGATTCCTCTG	900
	Dd	260	IleLeuGluIleMethHisLysGluAlaGlnAspIleLysPheThrGluGluIleProLeu	279
	QY	901	AAAACTCTGCACACAATGGCTTGTGTGAAGACTCATTTGGAAGAAAGACGAAATTTG	960
	Dd	280	LysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsnLeu	299
	QY	961	AAGAAAATTGAACATGAAACAGGACCAAGATAACAAATCTCATCTTTGACGATTGTAGC	1020
	Dd	300	LysLysIleGluGlnAspThrAspThrLysIleThrIleSerProLeuGlnGluLeuThr	319
	QY	1021	ATATACAAACCCGAAGAACCATCATGTGTAAGGGCACAGTTGAGCCCTGTGCCATGCT	1080
	Dd	320	LeuTyr-AsnProGluArgThrIleThrValLysGlyAsnValGluThrCysAlaLysAla	339
	QY	1081	GAGATAGAGATTATGAAGACTGCGGTGAGGCCCTTTGAAATGATATGCTGGCTGTTAAC	1140
	Dd	340	GluGluGluIleMetLysLysIleArgGluSer-Tyr-GluAsnAspIleAlaSerMetAsn	359
	QY	1141	CACCAAGCCAATCTGATCCAGGGTTGAACCTCAGCGCACTTGGCATCTTTTTCAACAGGA	1200
	Dd	360	LeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeuGlyLeuPhe-----	376
	QY	1201	CTGTCCGTGCTATCTCCACAGCAGGGGCCCGCGAGCTCCCCCGCTGCCCTACCCAC	1260
	Dd	377	-----ProProThrSer-----GlyMetProProThr-----	386
	QY	1261	CCCTTCACTACCCATCCGGATCTCTCCAGCGCTGTACCCCATCACAGTTTGGCCCG	1320
	Dd	387	-----SerGlyProProSerAlaMetThrPro-----Pro	396
	QY	1321	TTCGCCGATCATCACTCTTATCCAGACGAGAGATTGTGAATCTCTTCATCCCAACCCAG	1380
	Dd	397	TyrProGlnPheGluGln---SerGluThrGluThrValHisGlnPheIleProAlaLeu	415
	QY	1381	GCTGTGGCGGCATCATCGGAAAGAGGGGGCACACATCAACAGCTGGCGAGTTGCC	1440

```
Db 416 SerValGlyAlaIleGlyGlnGlnHisIleLysGlnLeuSerArgPheAla 435
QY 1441 GGAGCTCTATCAAGATTGCCCTCCGGAAGCCAGAGCTCAGCGAAGAGTGTCTATC 1500
Db 436 GlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLysValArgMetValIle 455
QY 1501 ATCCCGGGCCACCGGAGCCAGTTCAGGCCCGCAGGACCGATCTTTGGGAACTGAAA 1560
Db 456 IleThrGlyProGluGlnGlnPheLysAlaGlnGlyArgIleTyrGlyLysIleLys 475
QY 1561 GAGGAAACTTTTAAACCCCAAGAAAGTGAAGCTGGAGCCGATATCAGAGTGGCC 1620
Db 476 GluGluAsnPheValSerProLysGluGluValLysLeuGluAlaHisIleArgValPro 495
QY 1621 TCTTCCACAGCTGGCGGGTGAATTGGCAAGTGGCAAGCGTGAACGAACTGCAGAAC 1680
Db 496 SerPheAlaAlaGlyArgValIleGlyLysGlyGlyLysThrValAsnGluLeuGlnAsn 515
QY 1681 TTAACCACTGAGAGTCTATCGTCTGTCACCAAGCCAGATGAAATGAGGAAGTG 1740
Db 516 LeuSerSerAlaGluValValProArgAspGlnThrProAspGluAsnAspGlnVal 535
QY 1741 ATCGTCAGATTATGGGCACTTCTTCTAGCCAGACTGCACAGCGCAAGATCAGGAAA 1800
Db 536 ValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGlnArgLysIleGlnGlu 555
QY 1801 ATTGTACAAAGTGAAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1833
Db 556 IleLeuThrGlnValLysGlnHisGlnGlnGln 566
```

RESULT 8

```
US-09-606-421B-176
; Sequence 176, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 176
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-606-421B-176
```

```
Alignment Scores:
Pred. No.: 1.69e-164 Length: 579
Score: 1950.50 Matches: 386
Percent Similarity: 79.02% Conservative: 81
Best Local Similarity: 65.31% Mismatches: 95
Query Match: 31.87% Indels: 29
DB: 4 Gaps: 8
```

US-09-270-437D-6 (1-3412) x US-09-606-421B-176 (1-579)

```
QY 73 ATGACAGCTTTATCATCGGAACTGAGCCCGCCCTCAGCCCGACGACCTCCGCGAG 132
Db 1 MetAsnLysLeuTyrIleGlyAsnLeuSerGluAsnAlaAlaProSerAspLeuGluSer 20
QY 133 CTCTTTGGGACAGGAAGCTGCCCTTGGCGGACAGGTCCTGCTGAAGTCCGGCTAGGCC 192
```

```
Db 21 IlePheLysAspAlaLysIleProValSerGlyProPheLeuValLysThrGlyTyrAla 40
QY 193 TTCTGGGACTACCCGACGACAGAACTGGGCCATCCGCGCCATCGACACCTCTCTCGGGTAAA 252
Db 41 PheValAspCysProAspGluSerIrpAlaLeuLysAlaIleGluAlaLeuSerGlyLys 60
QY 253 GTGGAATTGCATGGGAAAATCATGGAAGTTGATTACTCAGTCTCTTAAAGAGCTAAGGAGC 312
Db 61 IleGluLeuHisGlyLysProIleGluValGluHisSerValProLysArgGlnArgIle 80
QY 313 AGGAAATTCAGATTCGAAACATCCCTCCTCAGCTGAGTGGAGGTGTGATGAGCTT 372
Db 81 ArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspSerLeu 100
QY 373 TTGGCTCAATATGGACAGTGCAGAAATGTGAAACAAGTCAACACAGACACACAGAACCCGC 432
Db 101 LeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAspSerGluThrAla 120
QY 433 GTTGTCAACGTCATATGCAACAGAGAAGCAAAATAGCCATCGAGAAAGCTAAGC 492
Db 121 ValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn 140
QY 493 GGGCATCAGTTTGAGAACTACTCTTCAAGATTTCTACATCCCGAGTGAAGAGGTGAGC 552
Db 141 GlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAspGluMetAlaAla 160
QY 553 TCCCTTCGCCCCCTCAG-----CGAGCCAGAGCT-----GGGACCACCTCTTCCCGG 600
Db 161 GlnGlnAsnProLeuGlnGlnProArgLysArgArgGlyLeuGlyGlnArgGlySerSer 180
QY 601 GAGCAAGGCCACGCCCTCGGGGCACCTTCTCAGGCCACAGACAGATTTGATTTCCCTCGCG 660
Db 181 ArgGlnGly--SerProGlySerValSerLysGlnLysProCysAspLeuProLeuArg 199
QY 661 ATCCTGCTCCCGCCCGGATTTGTTGGTCCCATCATCGGAAGAGGGGCTTGACCATAAAG 720
Db 200 LeuLeuValProThrGlnPheValGlyAlaIleGlyLysGlyLysGlyAlaThrIleArg 219
QY 721 AACATCACTAAGACAGACCCAGTCCCGGTAGATATCCATAGAAAGAGAGACTCTGGAGCT 780
Db 220 AsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGlyAla 239
QY 781 GCAGAGAAGCTGTACCATCCATCCACCCAGGAGGAGCTTCTGAAGATGCGCGATG 840
Db 240 AlaGluLysSerIleThrIleLeuSerThrProGluGlyThrSerAlaAlaCysLysSer 259
QY 841 ATTCTTGAATCATGCAAGAAGAGGAGGAGTGAACCAACTAGCCGAGAGATTCTCTGT 900
Db 260 IleLeuGluIleMetHisLysGluAlaGlnAspIleLysPheThrGluGluIleProLeu 279
QY 901 AAAATCTTGGCACACAATGGCTTGGTTGAAGACTGATTGGAAGAGGAGGAGGAGGAGGAG 960
Db 280 LysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGlyLysGlyArgGlnLeu 299
QY 961 AAGAAATTCGAACATGAACAGGAGGACCAAGATAACATCTCATCTTTGCGAGATTTCAGC 1020
Db 300 LysLysIleGluGlnAspThrAspThrLysIleThrLysSerProLeuGlnGluLeuThr 319
QY 1021 ATATACACCGGAAAGAACCATCAGTGAAGGGACAGCTTGAAGGCTGTGCGAGTCT 1080
Db 320 LeuTyrAsnProGluArgThrIleThrValLysGlyAsnValGluThrCysAlaLysAla 339
QY 1081 GAGATAGAGATTATGAAGAGCTGCTGAGGCGCTTTGAAATCATATGCTGGCTGTTTAAAC 1140
Db 340 GluGluGluIleMetLysLysIleArgGluSerTyrGluAsnAspIleAlaSerMetAsn 359
QY 1141 CAAACAAGCCAATCGATCCCGGGTTGAACCTCAGCGCACTTGGCATCTTTTCAACAGGA 1200
Db 360 LeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeuGlyLeuPhe----- 376
QY 1201 CTCTCCGTGCTATCTCCACAGCAGGCGGCCCGGAGCTCCCGCCCTGCCCCCTACCAC 1260
```



```

;
; ZIP: 10154
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
;
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
;
; APPLICATION NUMBER: PCT/US94/01782
; FILING DATE: 22-FEB-1994
;
; CLASSIFICATION:
;
; PRIOR APPLICATION DATA:
;
; APPLICATION NUMBER: U.S. 08/021,608
; FILING DATE: 22-FEB-1993
;
; ATTORNEY/AGENT INFORMATION:
; NAME: WILLIAM S. FEILER
;
; REGISTRATION NUMBER: 26,728
; REFERENCE/DOCKET NUMBER: 2026-4063PCT
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
;
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 643
; TYPE: Amino Acid
; STRANDEDNESS: Single
; TOPOLOGY: Unknown
;
; MOLECULE TYPE: Peptide/Protein
;
; HYPOTHETICAL: No
; ORIGINAL SOURCE:
; ORGANISM: Human
; STRAIN:
;
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
;
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE: HL60
; ORGANELLE:
; FEATURE:
;
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
;
; OTHER INFORMATION: Amino Acid 148 (Xaa) is Met or Ile
;
; PCT-US94-01782-10
;
; Alignment Scores:
; Pred. No.: 8.72e-13 Length: 643
; Score: 243.50 Matches: 145
; Percent Similarity: 35.48% Conservative: 81
; Best Local Similarity: 22.76% Mismatches: 241
; Query Match: 3.98% Indels: 170
; DB: 5 Gaps: 27
;
; US-09-270-437D-6 (1-3412) x PCT-US94-01782-10 (1-643)
;
; QY 336 CCCTCCCTCAGTCAGTGGAGGTGGTGGCTCAATATGGGACAGTGA 395
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; QY 10 ProSerGlySerAlaGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 22
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; QY 396 GAATGTGGAACAAGTCAACACAGACACAGACCGCGTGTGCAACGTCA 455
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; QY 23 GlyGlyGly-GlyValAsnAspAlaPheLysAspAlaLeuGlnArg 39
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; QY 456 AAGAGAAGAAAGAAAAATAGCCATGAGAGCTAAGCGGGCATCAGTTTGA 515
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; QY 39 gGlnIleAlaAlaLysIleGlyGlyAspAlaGlyThrSerLeuAsnSerAsn 59
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; QY 516 CTTCAGATTTCCATCCCGGATGAGAGGTGAGTCCCTTCGCCCTCAGCAGC 575
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; QY 59 YTyrglyGlyGlnLysArgProLeuGluAspGlyAspGlnProAspAlaLys 79
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; QY 576 CCAGCGTGGGACCACTCTTCCCGGAGCAGGCCACGCCCTGGGGCACTTCTCAGG 635

```

```

;
; Db 79 aProGlnAsnAspSerPheGlyThrGln-----LeuProMetHisGlnGln 96
; QY CAGACAGATTGATTTCCGCTGCGGATCCTCGTCCACCCAGTCTTGTGTGGCCATCAT 695
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; QY 96 nArgSerVal---MetThrGluGluTyrlsValProAspGlyMetValGlyPheIle 115
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; QY 696 CGGAAAGAGGGCTTGACCATATAAGACACCCAGTCCCGGGTAGATAT 755
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; QY 115 eGlyArgGlyGlyGluGlnIleSerArgIleGlnGlnLysSerGlyCysLysIleGln 135
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; QY 756 CCATAGAAAAGAGAACTCTCGAGCTGCAGAAAGCCGTCCACATCCATCCACCCAGA 815
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; QY 135 e---AlaProAspSerGlyGlyLeuProGluArgSerCys***LeuThrGlyThrPro 154
; QY 816 GGGGACTTCTGAAGCATGCCGATGATTCTTGAATCATGCAGAAA----- 861
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; QY 154 uSerValGlnSerAlaLysArgLeuLeuAspGlnIleValGlnLysGlyArgProAla 174
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; QY 862 -----GAGGCAGATGAGACCAACTAGCCGAAGAGATTCTCTGAAATCTTTGGC 911
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; QY 174 cGlyPheHisHisGlyAspGlyProGlyAsnAlaValGln-----GluIleMet 191
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; QY 912 ACACATGGCTTGGTGGAGACTGATTGGAAGAAGAGGAGCAAAATTTGAAGAAA 971
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; QY 191 eProAlaSerLysAlaGlyLeuValIleGlyLysGlyGlyGluThrIleLysGln 211
; QY 972 ACATGAAACAGAGGACCAAGATCAATCTCATCTTTCGAGGATTTGAGCATATCA 1031
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; QY 211 nGluArgAlaGlyValLysMetValMet-----IleGlnAsp-----GlyPr 225
; QY 1032 GGAAGAAGACCATCTCTGTGAAGGCGACAGTGTGGCCCTGTGCCAGTGTGAGAT 1091
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; QY 225 cGlnAsnThr-----GlyAlaAsp----- 231
; QY 1092 TATGAAGAAGCTCGTGCAGGCCCTTTGAAATGATATGCTGGCTGTAAACCAAGCC 1151
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; QY 232 ---LysProLeuArg-----IleThrGlyAspProTyrLysValGlnAlaLys 248
; QY 1152 TCTGATCCAGGGTTGAACCTCAGCGCACTGGCATCTTTTCAACAGAGCTGCGTGCT 1211
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; QY 248 uMetVal----- 254
; QY 1212 ATCTCCACAGAGGCGCGCGAGCTCCCGCGTCCCGCTGCCCTTACCACCCCTCACTAC 1271
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; QY 254 eArgAspGlnGlyGlyPheArgGluVal----- 263
; QY 1272 CCCTCCGGATATCTTCCAGCCTGTACCCCCATCACAGTTGGCGCGTCCCGCATCA 1331
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; QY 264 -----ArgAsnGluTyrGly----- 268
; QY 1332 TCACTCTTATCCAGACGAGGATGTGAATCTCTTCCATCCCAACCCAGGCTGGCGC 1391
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; QY 269 -SerArgIleGlyGlyAsnGluGlyIleAspValProIleProArgPheAlaValGly 288
; QY 1392 CATCATCGGAAGAGGGGCGACACATAACAGTGGCGAGATTGCGCGAGCCTCTAT 1451
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; QY 288 eValIleGlyArgAsnGlyGluMetIleLysLysIleGlnAsnAspAlaGlyValArg 308
; QY 1452 CAAGATTGCCCTCGGAAAGGCCAGCTCAGCGAAGAGGTGTCATCATCCCGGGCC 1511
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; QY 308 eGlnPheLysProAspAspGly---ThrThrProGluArgIleAlaGlnIleThrGlyPr 327
; QY 1512 ACCGGAA---GCCAGTTCAAGGCCAG----- 1536
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; QY 327 OProAspArgCysGlnHisAlaAlaGluIleIleThrAspLeuLeuArgSerValGln 347
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; QY 1537 -----GGACGGATCTTTGGGAAA----- 1554
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; QY 347 aGlyAsnProGlyGlyProGlyGlyArgGlyArgGlyGlyGlyGlyGlyGlyGly 367
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; QY 1555 -----CTGAAGAGGAGAAACTTCTTTAAACCCCAAGAGAAGT 1592
; Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db 367 pAsnMetGlyProProGlyGlyLeuGlnGluPheAsnPhelle----- 381
QY 1593 GNAGCTGAGCGCATATCAGAGTGCCTCTCCACAGCTGCGCGGGTGATGGCAAGG 1652
Db 382 -----ValProThrGlyLysThrGlyLeuIleGlyLysG 394
QY 1653 TGGCAAGACCGTGAACGAACCTTACCAGTGCAGAAAGTCATCGCTCGCTGA 1712
Db 394 yGlyGluThrIleLysSerGlnGlnSerGlyAlaArgIleGluLeuGlnArgAs 414
QY 1713 CCAAGCGCAGATGAAGAAATGAGAAAGTG---ATCGTCAGAATATCGGCAC- 1761
Db 414 nProProProAsnAlaAspProAsnMetLysLeuPheThrIleArgGlyThrProGlnG 434
QY 1762 ----TTCCTTGTAGCCAGACTGCACAGCGCAAGATCAGGAAATGTACACAGGTGA 1817
Db 434 nIleAspTyrAlaArgGlnLeuIleGluLysIleGlyProValAsnProLeuG 454
QY 1818 GCAGCAGGACGAGAAATACCTCAGGAGTGCCTCACAGCGCAGCAAGTGAGGTCCCA 1877
Db 454 yProProValProHisGlyProHisGlyVal-ProGlyProHisGlyProGlyProp 474
QY 1878 CAGGCACCGCAGCAAAACACGGATGATAGCCCTTCCAAACCTGCACAGATGACCA 1937
Db 474 roGly-ProGly-----ThrPro-----MetGlyPro 482
QY 1938 AACGCAGCGCAGCAGATCGGAGCAACCAAGACCATCTGAGGAATGAGAAGTCTGCGG 1997
Db 483 TyrAsnProAlaProTyrAsnProGlyPro----- 492
QY 1998 AGCGCGCAGGAGTCTGCGGAGG---CCCTGAGAACCCAGGGCGGAGGCGCGG 2054
Db 493 -----ProGlyProAlaProHisGlyProProAlaProTyrAlaProGlnGlyTrpGly 510
QY 2055 GAAGTGCAGCGAGTTTGCAGAA-----CCACGAGCGCGCGC 2093
Db 511 AsnAlaTyrProHisTrpGlnGlnGlnAlaProProAspProAla 525

RESULT 15

US-08-021-608D-8
Sequence 8, Application US/08021608D
Patent No. 5580760
GENERAL INFORMATION:
APPLICANT: LEVENS, DAVID L., DUNCAN,
APPLICANT: ROBERT C., AND AVIGAN, MARK I.
TITLE OF INVENTION: NOVEL FUSE BINDING
TITLE OF INVENTION: PROTEIN AND CDNA THEREFOR
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/021,608D
FILING DATE: 22-FEB-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: WILLIAM S. FEILER
REGISTRATION NUMBER: 26,728
REFERENCE/DOCKET NUMBER: 2026-4063
TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 590
TYPE: Amino Acid
STRANDEDNESS: Single
TOPOLOGY: Unknown
MOLECULE TYPE: Peptide/Protein
HYPOTHETICAL: No
ORIGINAL SOURCE: Human
CELL LINE: HL60
US-08-021-608D-8
Alignment Scores:
Pred. No.: 3,48e-12 Length: 590
Score: 236.50 Matches: 121
Percent Similarity: 35.93% Conservative: 68
Best Local Similarity: 23.00% Mismatches: 182
Query Match: 3.86% Indels: 155
DB: 23 Gaps: 23
US-09-270-437D-6 (1-3412) x US-08-021-608D-8 (1-590)
QY 667 GTCCCCACCCAGCTTTGTTGGTCCCATCATCGAAAGAGGGCTTGACCAATAAAGAACATC 726
Db 91 ValProAspGlyMetValGlyPheIleIleGlyArgGlyGlyGluGlnIleSerArgIle 110
QY 727 ACTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAGAGAACTCTGGAGCTGCAGAG 786
Db 111 GlnGlnGluSerGlyCysLysIleGlnIle---AlaProAspSerGlyGlyLeuProGlu 129
QY 787 AGCCTGTCCACCATCCATCCACCCAGAGGGGACTTCTGAAGCATGCCGCGCATGTTCTT 846
Db 130 ArgSerCysMetLeuThrGlyThrProGluSerValGlnSerAlaLysArgLeuLeuAsp 149
QY 847 GAAATCATGCAGAAA-----GAGGCAGATGAGACCAACTA 882
Db 150 GlnIleValGluLysGlyArgProAlaProGlyPheHisHisGlyAspGlyProGlyAsn 169
QY 883 GCCGAAGAGATTCCTCTGAAATCTTGGCACACAATGGCTTGGTTGGAAGACTGATGGA 942
Db 170 AlaValGln-----GluIleMetIleProAlaSerLysAlaGlyLeuValIleGly 186
QY 943 AAAGAAGCAGAAATTTGAGAAATTTGAACATGAACAGGACCCAGACATAACAATCTCA 1002
Db 187 LysGlyGlyGluThrIleLysGlnLeuGlnGluArgAlaGlyValLysMetValMet--- 205
QY 1003 TCTTTGCGAGGATTTGAGCATATACCAACCGGAAAGAACCATCACTGTGAAGGCGCACAGTT 1062
Db 206 ---IleGlnAsp-----GlyProGlnAsnThr----- 213
QY 1063 GAGGCTGTGCGAGTGTGAGATGATTAAGAAGCTGCTGGTGGAGCCCTTTGAAAT 1122
Db 214 -----GlyAlaAsp-----LysProLeuArg-----IleThrGly 223
QY 1123 GATATGCTGGCTGTTAAACCAACAGCAATCTGATCCAGGGTGTGAACCTCAGCGCACTT 1182
Db 224 AspProTyrLysValGlnGlnAlaLysGluMetVal----- 235
QY 1183 GGCATCTTTTCAACAGGACTGTCCGTGTCTATCTCCACCAGAGGGCCCGCGAGTCCC 1242
Db 236 -----LeuGluLeuIleArgAspGlnGlyGlyPheArgGluVal--- 248
QY 1243 CCCGCTGCCCCCTACACCCCTTCACTACCCACTCGGGTACTTCTCCAGCCTGTACCCC 1302
Db 248 ----- 248
QY 1303 CATCACCAGTTTGGCCCGTTCCCGCATCATCACTCTTATCCAGAGCAGGAGATGTGAAT 1362
Db 249 ArgAsnGluTyrGly-----SerArgIleGlyGlyAsnGluGlyIleAsp 263

```
QY 1363 CTCTTCATCCACCCAGGCTGTGGCGCCATCATCGGAGAAAGGGGGCACACATCAAA 1422
Db   :: ||||| ||||| ||||| ||||| ::||| ||||| ||||| |||||
264 ValProIleProArgPheAlaValGlyLeValIleGlyArgAsnGlyGluMetIleLys 283
QY 1423 CAGTGGCGAGATTCGGCGAGCCCTCTATCAAGATTGCCCTGGGAAGGCCAGACGTC 1482
Db   ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
284 LysIleGlnAsnAspAlaGlyValArgIleGlnPheLysProAspAspGly---ThrThr 302
QY 1483 AGCAAAAGGATGTCATCATCACGGGCCACCGGAA---GCCAGTTCAAGCCCGAG--- 1536
Db   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
303 ProGluArgIleAlaGlnIleThrGlyProProAspArgCysGlnHisAlaAlaGluIle 322
QY 1537 -----GGG 1539
Db   |||||
323 IleThrAspLeuLeuArgSerValGlnAlaGlyAsnProGlyGlyProGlyGlyGly 342
QY 1540 CGGATCTTTGGGAAA-----CTGAAAGAG 1563
Db   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
343 ArgGlyArgGlyArgGlyGlnGlyAsnTrpAsnMetGlyProProGlyGlyLeuGlnGlu 362
QY 1564 GAAACTTCTTTAAACCCCAAGAGAGAGTGAAGCTGGAAGCGCATATCAGAGTCCCTCT 1623
Db   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
363 PheAsnPheIle-----ValProThr 369
QY 1624 TCCACAGCTGCCGGGTGATTCGCAAGGTGGCAAGACCGTGAACGAACTGCAGAACTTA 1683
Db   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
370 GlyLysThrGlyLeuIleIleGlyLysGlyGlyGluThrIleLysSerIleSerGlnGln 389
QY 1684 ACCAGTCGCAAGATCGTCCCTCGTGACCAACCGCCAGATGAAATAGAGGAAGTG--- 1740
Db   ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
390 SerGlyAlaArgIleGluLeuGlnArgAsnProProAsnAlaAspProAsnMetLys 409
QY 1741 ATCCTCAGATATTCGGGCAC-----TTCCTTGTAGCCAGACTGCACAGCGC 1788
Db   ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
410 LeuPheThrIleArgGlyThrProGlnGlnIleAspTyrAlaArgGlnLeuIleGluGlu 429
QY 1789 AAGATCAGGGAAATTGTACACAGGTGAAGCAGCAGGAGCAGAAATACCTCAGGGAGTC 1848
Db   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
430 LysIleGlyGlyProValAsnProIleuGlyProProValProHisGlyProHisGlyVal 449
QY 1849 GCCTCAGCGCAGCAAGTGAGGCTCCACAGCGCAGCAGCAAAACACCGATGAATGTAG 1908
Db   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
450 -ProGlyProHisGlyProProGlyProProGly-ProGly----- 462
QY 1909 CCCTTCCACACCTGACAGATGAGACCAACCGCAGCCAGCCAGATCGGGAGCAACCAA 1968
Db   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
463 -----ThrPro-----MetGlyProTyrAsnProAlaProTyrAsnProGlyPro- 477
QY 1969 AGACCATCTGAGGAATGAGAAGTCTCGGAGCGCGCCAGGACTCTGCCGAGG---CCCT 2025
Db   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
478 -----ProGlyProAlaProHisGlyProp 486
QY 2026 GAGAACCCCGGGCCGAGGCGGGGAGGTTCAGCCAGGTTTGCCAGAA-----C 2079
Db   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
486 roAlaProTyrAlaProGlnGlyTrpGlyAsnAlaTyrProHisTrpGlnGlnGlnAlap 506
QY 2080 CACCGAGCCCGCC 2093
Db   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
506 roProAspProAla 510
```

Search completed: July 16, 2004, 10:47:06
Job time : 83 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: July 16, 2004, 10:33:46 ; Search time 318 Seconds
(without alignments)
6707.276 Million cell updates/sec

Title: US-09-270-437d-6
Perfect score: 6121
Sequence: 1 ggcagcgaggagcgaggga.....aaccttgaaaattttttttt 3412

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1285345 seqs, 312560633 residues

Total number of hits satisfying chosen parameters: 2570690

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US09270437/runat 16072004 113051 13449/app query.fasta_1.3591
-DB=Published Applications AA -QFMT=fastan -SUFFIX=rapb -MINMATCH=0.1
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsu62
-TRANS-human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct-THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HAPSIZ=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09270437 @CGN 1 1 250 @runat 16072004 113051 13449
-NCPU=6 -ICPU=3 -NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB ID	Description

1	3065.5	50.1	620	9	US-09-764-864-1116	Sequence 1116, Ap
2	3003	49.1	587	15	US-10-313-986-501	Sequence 501, App
3	2806.5	45.9	556	14	US-10-097-340-147	Sequence 147, App
4	2806.5	45.9	556	16	US-10-648-593-182	Sequence 182, App
5	2553	41.7	555	15	US-10-262-445-40	Sequence 40, Appl
6	2025.5	33.1	577	15	US-10-313-986-500	Sequence 500, App
7	2019.5	33.0	577	9	US-09-873-637-2	Sequence 2, Appli
8	1968.5	32.2	589	15	US-10-313-986-486	Sequence 486, App
9	1957.5	32.0	579	9	US-09-735-705-348	Sequence 348, App
10	1957.5	32.0	579	9	US-09-850-716A-348	Sequence 348, App
11	1957.5	32.0	579	9	US-09-897-778-348	Sequence 348, App
12	1957.5	32.0	579	9	US-09-897-778-446	Sequence 446, App
13	1957.5	32.0	579	9	US-09-897-778-449	Sequence 449, App
14	1957.5	32.0	579	12	US-10-007-700-348	Sequence 348, App
15	1957.5	32.0	579	12	US-10-007-700-446	Sequence 446, App
16	1957.5	32.0	579	12	US-10-007-700-449	Sequence 449, App
17	1957.5	32.0	579	14	US-10-117-982-348	Sequence 348, App
18	1957.5	32.0	579	14	US-10-117-982-446	Sequence 446, App
19	1957.5	32.0	579	14	US-10-117-982-449	Sequence 449, App
20	1957.5	32.0	579	14	US-10-117-982-480	Sequence 480, App
21	1957.5	32.0	579	15	US-10-313-986-348	Sequence 348, App
22	1957.5	32.0	579	15	US-10-313-986-446	Sequence 446, App
23	1957.5	32.0	579	15	US-10-313-986-449	Sequence 449, App
24	1957.5	32.0	579	15	US-10-313-986-480	Sequence 480, App
25	1952.5	31.9	586	9	US-09-850-716A-427	Sequence 427, App
26	1952.5	31.9	586	9	US-09-897-778-427	Sequence 427, App
27	1952.5	31.9	586	12	US-10-007-700-427	Sequence 427, App
28	1952.5	31.9	586	14	US-10-117-982-427	Sequence 427, App
29	1952.5	31.9	586	15	US-10-313-986-427	Sequence 427, App
30	1950.5	31.9	579	9	US-09-735-705-176	Sequence 176, App
31	1950.5	31.9	579	9	US-09-850-716A-176	Sequence 176, App
32	1950.5	31.9	579	9	US-09-897-778-176	Sequence 176, App
33	1950.5	31.9	579	10	US-09-466-396A-176	Sequence 176, App
34	1950.5	31.9	579	12	US-10-007-700-176	Sequence 176, App
35	1950.5	31.9	579	14	US-10-117-982-176	Sequence 176, App
36	1950.5	31.9	579	15	US-10-313-986-176	Sequence 176, App
37	1947.5	31.8	579	14	US-10-117-982-484	Sequence 484, App
38	1947.5	31.8	579	15	US-10-313-986-484	Sequence 484, App
39	1408.5	23.0	422	16	US-10-408-765A-2088	Sequence 2088, Ap
40	910	14.9	192	9	US-09-764-864-1117	Sequence 1117, Ap
41	898	14.7	261	9	US-09-764-864-1114	Sequence 1114, Ap
42	868	14.2	250	9	US-09-764-864-1532	Sequence 1532, Ap
43	803	13.1	171	9	US-09-764-864-1119	Sequence 1119, Ap
44	750	12.3	171	9	US-09-764-864-1536	Sequence 1536, Ap
45	357	5.8	81	14	US-10-117-982-476	Sequence 476, App

ALIGNMENTS

RESULT 1
US-09-764-864-1116
; Sequence 1116, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ23
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1116
; LENGTH: 620
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (533)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1116

Alignment Scores:

; CURRENT APPLICATION NUMBER: US/10/313,986
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 501
; LENGTH: 587
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-313-986-501

Alignment Scores:

Pred. No.: 8, 49e-228 Length: 587
Score: 3003.00 Matches: 587
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 49.06% Indels: 0
DB: 15 Gaps: 0

US-09-270-437D-6 (1-3412) x US-10-313-986-501 (1-587)

QY	73	ATGAACAAGCTTTACATCGGGAACCTGAGCCCGCGTCAACCGCGAGCACTCCGGCAG	132
Db	1	MetAsnLysLeuTyrlleGlyAsnLeuSerProAlaValThrAlaAspAspLeuArgGln	20
QY	133	CTCTTTGGGACAGGAGCTGCCCTCGCGGACAGTCTGCTGAAGTCGGCTACGCC	192
Db	21	LeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLeuLysSerGlyTyrAla	40
QY	193	TTCTGTGACTACCCCGACACAGAACTGGGCCATCGCGCCATCGAGACCCCTCTCGGTA	252
Db	41	PheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGluThrLeuSerGlyLys	60
QY	253	GTGGAATTGATGGGAAAATCATGGAAGTTGATTAATCTAAGTCTATAAAGAGCTAAG	312
Db	61	ValGluLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLysLeuArgSer	80
QY	313	AGGAAATTCAGATTCAAAATCCCTCCTCAGCTGAGTGGGAGTGTGGATGACCTT	372
Db	81	ArgLysIleGlnIleArgAsnIleProHisLeuGlnTrpGluValLeuAspGlyLeu	100
QY	373	TTGGCTCAATATGGGACAGTGGGAATGTGGAACAGTCAACACAGACACAGAAACCGCC	432
Db	101	LeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThrAspThrGluThrAla	120
QY	433	GTGTCAAGTTCATATGCAACAGAGAGAAAGCAAAAATAGCCATGGAGAAAGCTAAGC	492
Db	121	ValValAsnValThrTyrAlaThrArgGluGluAlaLysIleAlaMetGluLysLeuSer	140
QY	493	GGGCATCAGTTTGAGAACTACTCTTCAAGATTTCTACATCCCGGATGAGAGGTGAGC	552
Db	141	GlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIleProAspGluGluValSer	160
QY	553	TCCCTTCGCCCCCTCAGCGAGCCCGAGCTGGGACCACTCTTCCCGGAGCAAGGCCAC	612
Db	161	SerProSerProProGlnArgAlaGlnArgGlyAspHisSerSerArgGluGlnGlyHis	180
QY	613	GCCCTCGGGCACTTCTCAGCCAGACAGATTGATTTCCGCTCGGATCTCGTCCCTCCC	672
Db	181	AlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeuValPro	200
QY	673	ACCCAGTTGTGCTGCCATCATCGAAGAGGGCTTGACCATAAAGACATCACTTAAG	732
Db	201	ThrGlnPheValGlyAlaIleIleGlyLysGluGlyLeuThrIleLysAsnIleThrLys	220
QY	733	CAGACCCAGTCCCGGTAGATATCCATAGAAAAGAGAACTCTGGAGCTGCAGAGAGCCT	792
Db	221	GlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGlyAlaAlaGluLysPro	240
QY	793	GTCAACATCCATGCCACCCAGAGGGGACTTCTGAAGCATCCCGCATGATTTGAAATC	852
Db	241	ValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeuGluIle	260
QY	853	ATGCAGAAAGAGGCAGATGAGACCAACTAGCCGAGAGATTCCTCTGAAATATCTGGCA	912

Db	261	MetGlnLysGluAlaAspGluThrLysLeuAlaGluGluLeuProLeuLysIleLeuAla	280
QY	913	CACAATGGCTTGGTGAAGACTGATTGAAAAAGAGCGAGAATTTGAAGAAAATGAA	972
Db	281	HisAsnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLysIleGlu	300
QY	973	CATGAACAGGGACCAAGATAACAATCTCATCTTTGAGATTGAGCATATCAACCCG	1032
Db	301	HisGluThrGlyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsnPro	320
QY	1033	GAAAGAACCATCACTGTGAGGGCCAGCTTGGCCCTGTCAGTGTCTGAGATGAGATT	1092
Db	321	GluArgThrIleThrValLysGlyThrValGluAlaCysAlaSerAlaGluIleGluIle	340
QY	1093	ATGAAGAAGCTGCTGAGGCTTTGAAATGATATGCTGCTGTAAACCAACAAAGCCAAT	1152
Db	341	MetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaValAsnGlnAlaAsn	360
QY	1153	CTGATCCCGGGTGAACCTCAGGCACCTTGGGCACTCTTTTCAACAGGACTGTCTGCTA	1212
Db	361	LeuIleProGlyLeuAsnLeuSerAlaLeuGlyIlePheSerThrGlyLeuSerValLeu	380
QY	1213	TCTCCACCCAGCGCCCGCGAGCTCCCGCGCTGCCCGCTACCAACCCCTTCACTACC	1272
Db	381	SerProProAlaGlyProArgGlyAlaProProAlaAlaProTyrHisProPheThrThr	400
QY	1273	CACCTCCGGATPACTCTCCAGCTGTACCCCATCCAGCTTTGGCCCGTTCCTCCGCTCAT	1332
Db	401	HisSerGlyTyrPheSerSerLeuTyrProHisGlnPheGlyProPheProHisHis	420
QY	1333	CACCTTATCCAGACGAGAGATTGAAATCTCTCATCCCAACCCAGGCTGTGGGGCC	1392
Db	421	HisSerTyrProGluGlnGluIleValAsnLeuPheIleProThrGlnAlaValGlyAla	440
QY	1393	ATCATCGGAAGAGGGGCGACACATCAACAGCTGCGAGATTCCCGGAGGCTCTATC	1452
Db	441	IleIleGlyLysLysGlyAlaHisIleLysGlnLeuAlaArgPheAlaGlyAlaSerIle	460
QY	1453	AAGATTGCCCTCGGGAAGCCCGAGCTCAGCGAAAGGATGGTCAATCATCACCGGCCA	1512
Db	461	LysIleAlaProAlaGluGlyProAspValSerGluArgMetValIleIleThrGlyPro	480
QY	1513	CCGGAAGCCAGTTCAAGGCCCGGAGCGGATCTTTGGGAAACTGAAAGAGGAAACTTC	1572
Db	481	ProGluAlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluGluAsnPhe	500
QY	1573	TTTAAACCCCAAGAAAGTGAAGCTGGAAGCCCATATCAGAGTCCCTCTCCACAGCT	1632
Db	501	PheAsnProLysGluGluValLysLeuGluAlaHisIleArgValProSerSerThrAla	520
QY	1633	GGCCGGTGTATTGGCAAGGTGGCAGACCGGTGAACCACTGCAAGAACTTAAACAGTGCA	1692
Db	521	GlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsnLeuThrSerAla	540
QY	1693	GAGTCACTGCTGCTGTGACCAACCGCAGATGAAATCAGGAAGTATCGTCAGAAAT	1752
Db	541	GluValIleValProArgAspGlnThrProAspGluAsnGluGluValIleValArgIle	560
QY	1753	ATCGGGCACTTCTTTGCTAGCCAGCTGCACGCAAGATCAGGGAATTTGTACACAG	1812
Db	561	IleGlyHisPheAlaSerGlnThrAlaGlnArgLysIleArgGluIleValGlnGln	580
QY	1813	GTGAAGCAGCAGGAGCAGAAA 1833	
Db	581	ValLysGlnGlnGluGlnLys 587	

RESULT 3

US-10-097-340-147
; Sequence 147, Application US/10097340
; Publication No. US20030087250A1
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN

```
; APPLICANT: Manjula GANNAVARAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISSEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VREIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. EAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumel ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq For Windows Version 4.0
; SEQ ID NO 147
; LENGTH: 556
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-097-340-147

Alignment Scores:
Pred. No.: 2,66e-212 Length: 556
Score: 2806.50 Matches: 556
Percent Similarity: 92.82% Conservative: 0
Best Local Similarity: 92.82% Mismatches: 0
Query Match: 45.85% Indels: 43
DB: 14 Gaps: 1

US-09-270-437D-6 (1-3412) x US-10-097-340-147 (1-556)

QY 70 ATGATGAACAAGCTTTACATCGGGAACCTGAGCCCGCGGTCCACGGCGACGACCTCCGG 129
Db 1 MetMetAsnLysLeuTyrlleGlyAsnLeuSerProAlaValThrAlaAspLeuArg 20

QY 130 CAGCTCTTTGGGGACAGGAAGTGCCTCGTGGGGACAGGTCCTCTGAAGTCCGGCTAC 189
Db 21 GlnLeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLeuLysSerGlyTyr 40

QY 190 GCCTTCGTGGACTACCCGACACAGAACTGGGCGCATCCGCGGCATCGAGACCCCTCTCGG 249
Db 41 AlaPheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGluThrLeuSerGly 60

QY 250 AAAGTGAATTCGATGGGAAATCATGGAAGTTGATTACTCAGTCTCTAAAAGCTAAGG 309
Db 61 LysValGluLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLysLeuArg 80

QY 310 AGCAGGAAATTCAGATTCGAAACATCCCTCTCACTCGAGTGGGAGGTGTGGATGGA 369
Db 81 SerArgLysIleGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspGly 100

QY 370 CTTTGGCTCATATGGGACAGTGGAGATGTGGNACAAGTCAACACAGACACAGAAACC 429
```

```
Db 101 LeuLeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThrAspThrGluThr 120
QY 430 GCCGTTGTCAACGTCACATATGCAATGCAACAAGAGAGAAAGCAAAAATAGCCATGAGAGCTA 489
Db 121 AlaValValAsnValThrTyrAlaThrArgGluGluAlaLysIleAlaMetGluLysLeu 140
QY 490 AGCGGGCATCAGTTTGAGAACTACTCTTCAAGATTTCCTACATCCCGATCAAGAGGTG 549
Db 141 SerGlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIleProAspGluGluVal 160
QY 550 AGCTCCCTTCGCCCCCTCAGCGAGCCCGAGCGTGGGACCACTCTTCCCGGAGCAAGGC 609
Db 161 SerSerProSerProProGlnArgAlaGlnArgGlyAspHisSerSerArgGluGlnGly 180
QY 610 CACGCCCTTCGGGGCAGCTTCTCAGGCCACAGATTGATTTCCTCCCTCGGATCTCGTGC 669
Db 181 HisAlaProGlyGlyThrSerGluAlaArgGlnIleAspPheProLeuArgIleLeuVal 200
QY 670 CCCACCCAGTTTGTGGTCCCATCATCGAAAGAGGAGGCTTGACCATAAAGAACATCACT 729
Db 201 ProThrGlnPheValGlyAlaIleGlyLysGluGlyLeuThrIleLysAsnIleThr 220
QY 730 AAGCAGACCCAGTCCCGGTAGATATCCATAGAAAGAGAACTCTGGAGCTGCAGAGAG 789
Db 221 LysGlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGlyAlaAlaGluLys 240
QY 790 CTTGTACCATCCATCCGCCACCCAGAGGGAGCTTCTGAAGCATGCCGATGATCTTGA 849
Db 241 ProValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeuGlu 260
QY 850 ATCATGCAAGAGGAGGAGATGAGACCAACTAGCCGAGAGAGATTCCTCTGAAATCTTG 909
Db 261 IleMetGlnLysGluAlaAspGluThrLysLeuAlaGluGluIleProLeuLysIleLeu 280
QY 910 GCACAAATGCTTGGTTGGAAAGTGTGTTGAAAGAGGAGGAGGAGGAGGAGGAGGAGG 969
Db 281 AlaHisAsnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLysLysIle 300
QY 970 GAACATGAACACAGGACCAAGATTAACATCTCATCTTTGTCAGGATTTGAGATATACAA 1029
Db 301 GluHisGluThrGlyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsn 320
QY 1030 CCGGAAAGAACCATCACTGTCGAAGGACAGTTCAGGCTGTGCCAGGCTGAGATAGAG 1089
Db 321 ProGluArgThrIleThrValLysGlyThrValGluAlaCysAlaSerAlaGluIleGlu 340
QY 1090 ATTATGAAGAGCTGCGTGAGGCTTTGAAAATGATATGCTGGCTGTGTTAACCAAGCC 1149
Db 341 IleMetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaValAsn----- 357
QY 1150 AATCTGATCCAGGGTTGAACCTCAGCGCATCTGGCATCTTTTCAACAGGACTGTCCGTG 1209
Db 357 ----- 357
QY 1210 CTATCTCCACGAGGCGCGCGGAGCTCCCCCGCTGCCCCCTACCAACCCCTTCACT 1269
Db 357 ----- 357
QY 1270 ACCCACTCCGATATCTTCCAGCCTGTACCCCATCACAGTTTGGCCGTTCCCGCAT 1329
Db 358 ThrHisSerGlyTyrPheSerSerLeuTyrProHisGlnPheGlyProPheProHis 377
QY 1330 CATCACTCTTATCCAGAGCAGAGATTTGTGAATCTCTTCATCCCAACCCAGCGCTGTGGC 1389
Db 378 HisHisSerTyrProGluGlnGluIleValAsnLeuPheIleProThrGlnAlaValGly 397
QY 1390 GCCATCATCGGAAAGAGGGGGCACACATCAAAAGCTGGCGAGATTCGCGGAGCTCT 1449
Db 398 AlaIleIleGlyLysLysGlyAlaHisIleLysGlnLeuAlaArgPheAlaGlyAlaSer 417
QY 1450 ATCAAGATTGCCCCCTGGGGAAGGCCAGACCTCAGCGAAAGGATGGTGCATCATCCCGG 1509
```

Db 418 IleLysIleAlaProAlaGluGlyProAspValSerGluArgMetValIleIleThrGly 437
QY 1510 CCACGGAGCCAGTTCAAGCCACAGGACGGATCTTTGGGAAACTGAAAGAGGAAAC 1569
Db 438 ProProGluAlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluGluAsn 457
QY 1570 TTTCTTAACCCCAAGAGAGTGAAGCTGGAAGCGCATATCAGAGTCCCTCTTCCACA 1629
Db 458 PhePheAsnProLysGluGluValLysLeuGluAlaHisIleArgValProSerSerThr 477
QY 1630 GCTGCCGGGTGATTGGCAAGGTGGCAAGACCGTGAACGAAGTGCAGAACTTAACCAAGT 1689
Db 478 AlaGlyArgValIleGlyLysGlyGlyLysThrValAsnGluLeuGlnAsnLeuThrSer 497
QY 1690 GCAGAGTCATCGTCCCTCGTACCACCAAGCCAGATGAAATGAGGAAGTGAATCGTCA 1749
Db 498 AlaGluValIleValProArgAspGlnThrProAspGluAsnGluGluValIleValArg 517
QY 1750 ATTATCGGGCACTTCTTTGCTAGCCAGACTGCACAGCGCAAGATCAGGGAATTTGACAA 1809
Db 518 IleIleGlyHisPhePheAlaSerGlnThrAlaGlnArgLysIleArgGluIleValGln 537
QY 1810 CAGTGAAGCGAGGACAGAAATACCTCAGGAGTGCCTCAGCGCAGCAAG 1866
Db 538 GlnValLysGlnGlnGluGlnLysTyrProGlnGlyValAlaSerGlnArgSerLys 556
RESULT 4
US-10-648-593-182
; Sequence 182, Application US/10648593
; Publication No. US20040106132A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT
; TITLE OF INVENTION: INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
; FILE OF INVENTION: PROTEIN TYROSINE KINASE PATHWAYS IN BREAST CELLS
; FILE REFERENCE: D0273 NP
; CURRENT APPLICATION NUMBER: US/10/648,593
; PRIOR FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: 60/406,385
; PRIOR FILING DATE: 2002-08-27
; NUMBER OF SEQ ID NOS: 557
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 182
; LENGTH: 556
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-648-593-182
Alignment Scores:
Pred. No.: 2,66e-212 Length: 556
Score: 2806.50 Matches: 556
Percent Similarity: 92.82% Conservative: 0
Best Local Similarity: 92.82% Mismatches: 0
Query Match: 45.85% Indels: 43
DB: 16 Gaps: 1
US-09-270-437D-6 (1-3412) x US-10-648-593-182 (1-556)
QY 70 ATGATGAACAGCTTTACATCGGAACCTGAGCCCGCGCTACCGCGCAGCAGCTCCGG 129
Db 1 MetMetAsnLysLeuTyrIleGlyAsnLeuSerProAlaValThrAlaAspLeuArg 20
QY 130 CAGCTCTTTGGGACAGGAAGCTCCCTGCGGGGACAGGTCCTGCTGAAGTCGGGTAC 189
Db 21 GlnLeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLeuLysSerGlyTyr 40
QY 190 GCCTTCGTGGACTACCCGACAGAACTGGGCCATCGCGCCATCGAGACCTCTCGGGT 249
Db 41 AlaPheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGluThrLeuSerGly 60
QY 250 AAATGGAATTCATGGGAAATCATGGAAGTGTATTACTCAGTCTCTAAAGCTAAGG 309
Db 61 LysValGluLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLysLeuArg 80

QY 310 AGCAGGAAAAATTCAGATTTCGAAACATCCCTCTCACTGAGTGGAGGTGTGGATGGA 369
Db 81 SerArgLysIleGlnIleArgAsnIleProHisLeuGlnTrpGluValLeuAspGly 100
QY 370 CTTTGGCTCAATATGGGACAGTGGAGNAATGTGAAACAAGTCAACACAGACACAGAAACC 429
Db 101 LeuLeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThrAspThrGluThr 120
QY 430 GCCGTGTGCAACATATGCAACAGAGAAAGCAAGCAAAATAGCCATGGAGAAGCTA 489
Db 121 AlaValValAsnValThrTyrAlaThrArgGluAlaLysIleAlaMetGluLysLeu 140
QY 490 AGCGGCAATCAGTTTGAGAACTACTCTTCAAGATTCTCTACATCCCGATGAAGAGGTG 549
Db 141 SerGlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIleProAspGluGluVal 160
QY 550 AGCTCCCTTCGCCCCCTCAGCGAGCCAGCGTGGGACCACTCTTCCGGGAGCAAGGC 609
Db 161 SerSerProSerProGlnArgAlaGlnArgGlyAspHisSerSerArgGluGlnGly 180
QY 610 CACGCCCTCGGGGCACTTCTCAGGCCAGACAGATTGATTTCCGCTCGCGATCTCGTC 669
Db 181 HisAlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeuVal 200
QY 670 CCCACCCAGTTTGTGTGCCATCATCGAAAGAGGGCTTGACCATAAAGAACATCACT 729
Db 201 ProThrGlnPheValGlyAlaIleGlyLysGluGlyLeuThrIleLysAsnIleThr 220
QY 730 AAGCAGACCCAGTCCCGGTAGATATCCATAGAAAGAGACTCTGGAGTGCAGAGNAG 789
Db 221 LysGlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGlyValAlaGluLys 240
QY 790 CCTCTCACATCCATGCCACCCAGAGGGACTTCTGAAGCATCCGCGATCATTTCTGAA 849
Db 241 ProValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeuGlu 260
QY 850 ATCATGAGAAAGAGGAGATGAGACCAACTAGCCGAGAGATTCTCTGAAATCTTG 909
Db 261 IleMetGlnLysGluAlaAspGluThrLysLeuAlaGluGluIleProLeuLysIleLeu 280
QY 910 GCACACAATGCTTGTGGTGAAGACTGATTGGAAAGAGGACAGAAATTTGAAGAAATT 969
Db 281 AlaHisAsnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLysLysIle 300
QY 970 GAACATGAAACAGGACCAAGATAACAATCTCATCTTTTGGAGGATTTGAGCATATAAAC 1029
Db 301 GluHisGluThrGlyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsn 320
QY 1030 CCGGAAAGAACCATCACTGTAAGGGCAGATTTGAGGCTGTGCGAGTGTGAGATAGAG 1089
Db 321 ProGluArgThrIleThrValLysGlyThrValGluAlaCysAlaSerAlaGluIleGlu 340
QY 1090 ATTATGAAGAGTGCCTGAGGCTTTGAAATCATATGCTGCTGTATACCAACAAAGCC 1149
Db 341 IleMetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaValAsn----- 357
QY 1150 AATCTGATCCAGGGTTGAACCTCAGCGCACTTGGCATCTTTTCAACAGGACTGTCG 1209
Db 357 ----- 357
QY 1210 CTATCTCCACAGAGGGCCCCGGGAGCTCCCCCGCTGCCCTTACCACCCCTTCACT 1269
Db 357 ----- 357
QY 1270 ACCCACTCCGATACTTCTCCAGCTGTACCCCATCCACAGTTTGGCCGCTTCCCGCAT 1329
Db 358 ThrHisSerGlyTyrPheSerSerLeuTyrProHisGlnPheGlyProPheProHis 377
QY 1330 CATCACTCTTATCCAGACGAGGAGATTGTGAATCTCTTCATCCCAACCCAGGCTGTGGGC 1389
Db 378 HisHisSerTyrProGluGlnGluIleValAsnLeuPheIleProThrGlnAlaValGly 397

```
QY 1390 GCCATCATCGGAAGAGGGGSCACACATCAACAGCTGGCAGATTCCGCGGAGCCTCT 1449
Db 398 AlaIleIleGlyLysGlyAlaHisIleLysGlnLeuAlaArgPheAlaGlyAlaSer 417
QY 1450 ATCAAGATTGCCCTCGGGAAGGCCAGACGTCAGCGAAAGGATGTCATCATCCCGG 1509
Db 418 IleLysIleAlaProAlaGluGlyProAspValSerGluArgMetValIleIleThrGly 437
QY 1510 CCACCGAGGCCAGTCAAGCCCGCAGGACCGATCTTTGGGAAACTGAAAGAGAAAC 1569
Db 438 ProProGluAlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluGluAsn 457
QY 1570 TTTCTTAAACCCAAAGAGAGTGAAGCTGGAAGCGCATATCAGAGTCGCCCTTCCACA 1629
Db 458 PhePheAsnProLysGluGluValLysLeuGluAlaHisIleArgValProSerSerThr 477
QY 1630 GCTGCGCGGCTGATTGGCAAGAGTGGCAAGACCGTGAACGAACTGCAGAACTTAACCACT 1689
Db 478 AlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsnLeuThrSer 497
QY 1690 GCAGAAGTCATCGTGCCTCGTGCACCAAGCCAGATGAAATCAGGAAGTGTATCGTCAGA 1749
Db 498 AlaGluValIleValProArgAspGlnThrProAspGluAsnGluValIleValArg 517
QY 1750 ATTATCGGCACTTCTTTCTAGCCAGACTGCACAGCGCAAGATCAGGGGAAATTTGACAA 1809
Db 518 IleIleGlyHisPhePheAlaSerGlnThrAlaGlnArgLysIleArgGluIleValGln 537
QY 1810 CAGGTGAAGCAGCAGGAGCAGAAATACCTCTAGGAGTGCCTCTCAGCGCAGCAAG 1866
Db 538 GlnValLysGlnGlnGluGlnLysTyrProGlnGlyValAlaSerGlnArgSerLys 556

RESULT 5
US-10-262-445-40
; Sequence 40, Application US/10262445
; Publication No. US20040014058A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John
; APPLICANT: Burgess, Catherine
; APPLICANT: Catterton, Elina
; APPLICANT: Chant, John
; APPLICANT: Chaudhuri, Amitabha
; APPLICANT: Edinger, Shlomit
; APPLICANT: Gerlach, Valerie
; APPLICANT: Giot, Loic
; APPLICANT: Gorman, Linda
; APPLICANT: Guo, Xiaojia
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Mezes, Peter
; APPLICANT: Millet, Isabelle
; APPLICANT: Ooi, Chean Eng
; APPLICANT: Patturajan, Meera
; APPLICANT: Rieger, Daniel
; APPLICANT: Spytek, Kimberly
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Zhong, Haihong
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS
; FILE OF INVENTION: THE SAME
; FILE REFERENCE: 21402-462D
; CURRENT APPLICATION NUMBER: US/10/262,445
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/327,454
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,849
; PRIOR FILING DATE: 2001-10-12
```

```
; PRIOR APPLICATION NUMBER: 60/329,414
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 60/330,142
; PRIOR FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: 60/341,058
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: 60/343,629
; PRIOR FILING DATE: 2001-10-24
; PRIOR APPLICATION NUMBER: 60/349,575
; PRIOR FILING DATE: 2001-10-29
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 40
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-262-445-40

Alignment Scores:
Pred. No.: 2,68e-192 Length: 555
Score: 2553.00 Matches: 507
Percent Similarity: 87.81% Conservative: 19
Best Local Similarity: 84.64% Mismatches: 29
Query Match: 41.71% Indels: 44
DB: 15 Gaps: 2

US-09-270-437d-6 (1-3412) x US-10-262-445-40 (1-555)
QY 70 ATGATGAACAAGCTTTTACATCGGGAACCTGAGCCCGCGCTCACCGCCGACGACCTCCGG 129
Db 1 MetMetAsnLysLeuPheIleGlyAsnLeuSerProAlaValThrAlaGluAspLeuArg 20
QY 130 CAGCTCTTTGGGGACAGGAAGCTGCCCTCGCGGGACAGGTCTCTGCTGAAGTCCGGCTAC 189
Db 21 GlnLeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLysSerArgTyr 40
QY 190 GCCTTCGTGGACTACCCGACCCAGAACCTGGGCCATCCGCCCATCGAGACCCCTCTCGGCT 249
Db 41 AlaPheValAspTyrProAspGlnAsnTrpAlaIleArgThrIleGluThrLeuSerGly 60
QY 250 AAAAGTGAATTCATGCGGAAATCATGGAAGTTGATTACTCAGTCTCTAAAAAGCTAAGG 309
Db 61 GlnValGluLeuHisGlyLysIleMetGluValAspTyrSerValSerIleLysLeuArg 80
QY 310 AGCAGGAAATTCAGATTTCGAAACATCCCTCTCACCTGCGAGTGGAGGTGTGGATGGA 369
Db 81 SerArgAsnIleProIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspGly 100
QY 370 CTTTGGCTCAATATGGGACAGTGGAGATGTGGAACTGGAACAAAGTCAACACACACAGAAC 429
Db 101 LeuLeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThrAspThrGluThr 120
QY 430 GCCGTTGTCAACGTCAATATGCAATATGCAACAAGAGAAAGCAAAAATAGCCATGAGAGCTA 489
Db 121 AlaValValAsnValThrTyrAlaThrLysGluGluValLysIleAlaMetLysLysLeu 140
QY 490 ACGGGGCATCAGTTTGAGNACTACTCTTCAGATTCTTACATCCCGATCCGATCAAGGGTG 549
Db 141 SerGlyHisGlnPheGluAsnHisTyrPheLysIleSerTyrIleProAspAspGluVal 160
QY 550 AGCTCCCTTCGCGCCCTCAGCGAGCCCGAGCTGGGACCACTCTTCCCGGAGCAAGGC 609
Db 161 SerCysProSerProProGlnArgAlaGlnArgGlyAspHisSerSerTrpGluGlnGly 180
QY 610 CACGCCCTCGGGGCATCTTCTCAGGCCAGACAGATTGATTCCTCGCTCGGATCCTGGTC 669
Db 181 GlnAlaProGlyGlySerSerGlnAlaArgGlnIleAspPheProLeuArgValLeuPhe 200
QY 670 CCCACCCAGTTTGTGGTCCCATCATCGGAAAGGGGCTTCACCATTAAGAACATCACT 729
Db 201 ProThrGlnPheValGlyAlaIleIleGlyLysGluGlyLeuThrIleLysAsnIleThr 220
```

```
QY 730 AAGCAGACCCAGTCCGGGTAGATATCCATAGAAAGAGAACTCGAGCTCGAGGAG 789
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 221 LysGlnSerArgSerArgValAspIleIleArgGlnGluAsnSerArgAlaGluLys 240
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 790 CTTGTCCATCCATCCACCCAGAGGGGACTTCTGAAGCATGCCGATGATCTTGAA 849
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 241 ProValThrMetHisAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeuGlu 260
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 850 ATCATGCCAAGAGAGCAGATGAGACCAACTAGCGGAAGATTCCTCTGAAATCTTG 909
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 261 IleMetGlnLysGluAlaAspGluAlaLysLeuAlaGluIleProLeuLysIleLeu 280
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 910 GCACAATAGCTGCTGGTGGAGACTGATTGCAAGAGAGGAGCAAAATTTGAAGAAAT 969
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 281 AlaHisAsnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLysAsn 300
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 970 GAACATGAAACAGGACCAAGATAACAATCTCTATCTTTGCGAGATTGAGCATATACAAC 1029
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 301 GluHisGluThrGlyThrLysIleThrIleSerSerSerGlnAspLeuSerIleIleThrAsn 320
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 1030 CCGGAAGAACCATCACTGTGAGGCGACAGTTGAGGCGCTGGCAGTGTGAGATAGAG 1089
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 321 ProGluArgThrIleThrValLysGlyThrValGluValCysAlaSerAlaGluIleGlu 340
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 1090 ATTATGAAGAGCTCGGTGAGGCTTTGAAATGATATGCTGGCTGTTAAACCAAGCC 1149
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 341 IleMetLysLysLeuArgGluAlaPheGluAsnAspThrLeuThrValAsn----- 357
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 1150 AATCTGATCCCGAGGTTGAACCTCAGCGCACTTTGGCATCTTTTCAACAGGACTGTCCGTG 1209
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 357 ----- 357
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 1210 CTATCTCCACGAGGCGCGCGAGCTCCCGCGTCCCGCTGCGCCCTACACCCCTTCACT 1269
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 357 ----- 357
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 1270 ACCCACTCCGGATACTTCTCCAGCGCTGTACCCCATCACTGTTGGCCCGTCCCGCAT 1329
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 358 ThrHisPheGlyTyrPheSerSerLeuTyrProHisArgGlnPheGlyProPheProHis 377
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 1330 CATCACTCTTATCCAGCAGGAGATTGTGAATCTCTTTCATCCCAACCCAGCGCTGTGGCC 1389
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 378 HisSerTyrProGluGlnGluIleValAsnLeuPheIleProThrGlnGlyValGly 397
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 1390 GCGATCATCGGAAGAGGGGACACATCAACAGCTGGCGAGATTCCCGAGCTCT 1449
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 398 AlaIleIleGlyLysLysGlyAlaHisIleLysGlnLeuAlaArgPheValGlyAlaSer 417
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 1450 ATCAAGATTGCCCTCGCGAAGGCCAGACGCTCAGCGAAAGGATGCTCATCATCACCGG 1509
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 418 IleLysIleAlaProAlaArgSerPro---LeuArgGlnArgLysValIleIleThrTrp 436
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 1510 CCACGGAGCCCGAGTTCAAGCCCGAGGCGGATCTTTGGGAAACTGAAAGAGGAAAC 1569
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 437 ProProGluSerGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluAsn 456
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 1570 TTTCTTAACCCCAAGAGAGAGTGAAGCTGAAGCGGATATCAGAGTGCCCTCTTCCACA 1629
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 457 PhePheAsnProLysGluAspValLysLeuGluThrHisIleArgValProSerSerThr 476
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 1630 GCTGGCGGGTGATTGGCAAGGTGGCAAGACCGTGAACGACTGCAAGCTTAACAGT 1689
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 477 AlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsnLeuIleSer 496
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 1690 GCAGAAGTCATCGTCTCGTGACCAACAGCCAGATGAAATAGAGGAGTGCCTCAGA 1749
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 497 AlaGluValIleValProArgAspGlnThrProAspGluAsnGluGluMetIleValArg 516
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 1750 ATTATCGGCACTCTTCTGCTAGCAGACTCCACAGCGCAAGATCAGGAAATTTGACAA 1809
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 517 IleIleGlyHisPhePheAlaSerGlnThrAlaGlnArgLysIleArgGluIleValGln 536
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 1810 CAGGTGAAGCAGCAGGAGCAGAAATACCTTCAGGGAGTCCGCTCACAGCGCAGCAAG 1866
```

```
Db      537 GlnValLysGlnGlnGluGlnLysTyrProGlnGlyValAlaSerGlnArgSerLys 555
RESULT 6
US-10-313-986-500
; Sequence 500, Application US/10313986
; Publication No. US20030236209A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: McNabb, Andria
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C19
; CURRENT APPLICATION NUMBER: US/10/313,986
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 500
; LENGTH: 577
; TYPE: PXT
; ORGANISM: Homo sapiens
US-10-313-986-500
Alignment Scores:
Pred. No.:      1,15e-150      Length:      577
Score:          2025.50        Matches:     396
Percent Similarity: 78.94%      Conservative: 80
Best Local Similarity: 65.67%    Mismatches:  96
Query Match:    33.09%         Indels:     31
DB:             15            Gaps:       4
US-09-270-437D-6 (1-3412) x US-10-313-986-500 (1-577)
```

```
QY 73 ATGAAACAGCTTTTACATCGGGAACCTGAGCCCCCGCTCACCGCGACGACCTCCGGCAG 132
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 1 MetAsnLysLeuTyrIleGlyAsnLeuAsnGluSerValThrProAlaAspLeuLys 20
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 133 CTCTTTGGGACAGGAAGCTCCCTCGCGGAGAGTCTCTGTGAAGTCCGGTACGCC 192
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 21 ValPheAlaGluHisLysIleSerTyrSerGlyGlnPheLeuValLysSerGlyTyrAla 40
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 193 TTTCTGTGACTACCCCGACAGCAACTGGGCCATCGGCCATCGAGACCTCTCGGTAAA 252
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 41 PheValAspCysProAspGluHisTrpAlaMetLysAlaIleGluThrPheSerGlyLys 60
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 253 GTGAATTGATGGGAAATCATGGAAGTTGATTACTCAGTCTCTAAAAGCTAAGGAGC 312
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 61 ValGluLeuGlnGlyLysArgLeuGluIleGluHisSerValProLysLysGlnArgSer 80
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 313 AGGAAATTCAGATTCCGAACATCCCTCCTCACCTGAGTGGAGGTGTTGGATGGACTT 372
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 81 ArgLysIleGlnIleArgAsnIleProGlnLeuArgTrpGluValLeuAspSerLeu 100
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 373 TTGCTCAATATGGACAGTGGAGAATGTGAAACAAGTCAACACAGACACAGAACCGCC 432
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 101 LeuAlaGlnTyrGlyThrValGluAsnCysGluGlnValAsnThrGluSerGluThrAla 120
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 433 GTTGTCAAGTTCATATGCAACAAAGAGAAGAAATAGCCATGGAGAGCTAAGC 492
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 121 ValValAsnValThrTyrSerAsnArgGluGlnThrArgGlnAlaIleMetLysLeuAsn 140
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 493 GGGCATCAGTTTTCAGACTACTCTCTCAAGATTCTTACATCCCGGATGAAGGTGAGC 552
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 141 GlyHisGlnLeuGluAsnHisAlaLeuLysValSerTyrIleProAspGluGlnIleAla 160
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 553 TCCCTTCGCCCCCTCAGCAGAGCCCGAGCTGGGAGCCACTCTTCCCGGAGGAGGCCAC 612
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 161 -----GlnGlyProGluAsnGlyArgGlyGlyPheGlySerArgGlyGlnProArg 178
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 613 -----GCCCTCGGGGCACTTCTCAGGCCACAGACAGATTGATTTCCCGCTG 657
```

```
Db 179 GlnGlySerProValAlaAlaGlyAlaProAlaLysGlnGlnValAspIleProLeu 198
QY 658 CGGATCTGGTCCCAACCCAGTTTGTGGTCCATCATCGGAAAGAGGGCTTGACCAT 717
Db 199 ArgLeuLeuValProThrGlnTrValGlyAlaIleIleGlyLysGluGlyAlaThrIle 218
QY 718 AAGAATCATCAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAAGAGAACTCTGGA 777
Db 219 ArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGly 238
QY 778 GCTGAGAGAACCTGTCTACCATCCATGCCACCCAGAGGGGACTTCTGAAGCATGCCGC 837
Db 239 AlaAlaGluLysAlaIleSerValHisSerThrProGluGlyCysSerSerAlaCysLys 258
QY 838 ATGATCTTGAATCATGCAAGAAAGAGGAGATAGACCAAACTAGCCGAAGATTCCT 897
Db 259 MetIleLeuGluIleMetHisLysGluAlaLysAspThrLysThrAlaAspGluValPro 278
QY 898 CTGAAATCTTGGCACACAATGGCTTGGTTGGAAGACTGATTGGAAAAGAGCAGAAAT 957
Db 279 LeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsn 298
QY 958 TTGAAGAAATTTGAACATAAGACAGGACCAAGATAACAATCTCATTTTGAGAGTTTG 1017
Db 299 LeuLysLysValGluGlnAspThrGluThrLysIleThrIleSerSerLeuGlnAspLeu 318
QY 1018 AGCATATACACCCGAAAGAACCATCACTGTGAGGGCACAGTTGAGGCCCTGTGCCAGT 1077
Db 319 ThrLeuTyAsnProGluArgThrIleThrValLysGlyAlaIleGluAsnCysArg 338
QY 1078 GCTGACATAGAGATTGAAGAGCTGCGTAGGCTTTGAAATCATATGCTGGCTGT 1137
Db 339 AlaGluGlnGluIleMetLysLysValArgGluAlaTyArgLysAspValAlaAlaMet 358
QY 1138 AACCAACAGCCCAATCTGATCCAGGGTTGAACCTCAGCGCACTTGGCATTTTCAACA 1197
Db 359 SerLeuGlnSerHisLeuIleProGlyLeuAsnLeuAlaAlaValGlyLeuPheProAla 378
QY 1198 GGAATGTCGTGTATCTCCACAGCAGGGCCCCGGAGCTCCCCCGCTGCCCTAC 1257
Db 379 SerSerSerAlaValProPro-----ProSerSerValThrGlyAlaAlaProTy 396
QY 1258 CACCCCTTCACTACCCACTCCGGATCTTCTCCAGCCTGTATCCCCCATCACCAGTTTGGC 1317
Db 397 SerSerPheMet----- 400
QY 1318 CGGTTCCCGCATCATCACTCTTATCCAGACGAGAGATTGTGAATCTCTTCATCCCAAC 1377
Db 401 -----GlnAlaProGluGlnGluMetValGlnValPheIleProAla 414
QY 1378 CAGGCTGTGGCGCCATCATCGGGAAGAGGGGCGCACATCAACAGCTGGCGAGATTC 1437
Db 415 GlnAlaValGlyAlaIleIleGlyLysLysGlyGlnHisLysLysGlnLeuSerArgPhe 434
QY 1438 GCCGAGACCTCTCAAGATTGCCCCCTCGGAGAGCCAGACCTCAGCAAAAGAGTGTTC 1497
Db 435 AlaSerAlaSerIleLysIleAlaProProGluThrProAspSerLysValArgMetVal 454
QY 1498 ATCATACCGGGCCACCGGAAGCCAGTTCAAGGCCCAGGACGAGATCTTTGGGAACCTG 1557
Db 455 IleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTyGlyLysLeu 474
QY 1558 AAGAGGAAACCTCTTTTAAACCCCAAGAGAGAGTGAAGCTGAAGCGCATATCAGAGTG 1617
Db 475 LysGluGluAsnPhePheGlyProLysGluGluValLysLeuGluThrHisIleArgVal 494
QY 1618 CCCTCTTCCACAGCTGCCGGGTGATTGGCAAGAGTGGCAAGACCGTGAACGAACTGCAG 1677
Db 495 ProAlaSerAlaAlaGlyArgValIleGlyLysGlyGlyLysThrValAsnGluLeuGln 514
QY 1678 AACTTAACAGTACAGAGTCACTGCTCCCTCGTACCAACCGCCAGTGAATGAGGAA 1737
```

```
Db 515 AsnLeuThrAlaAlaGluValValProArgAspGlnThrProAspGluAsnAspGln 534
QY 1738 GTGATCGTCAGAAATATCGGCACATCTTTGTAGCCAGACTGCACAGCGCAAGATCAG 1797
Db 535 ValIleValLysIleIleGlyHisPheTyAlaSerGlnMetAlaGlnArgLysIleArg 554
QY 1798 GAAATTTACACAGCTGAAGCAGCAGGACGAGAAATACCTTCAGGGAGTCCCTCACAG 1857
Db 555 AspIleLeuAlaGlnValLysGlnHisGlnLysGlyGlnSerAsnGlnAlaGlnAla 574
QY 1858 CCGACGCAAG 1866
Db 575 ArgArgLys 577
RESULT 7
US-09-873-637-2
; Sequence 2, Application US/09873637
; Patent No. US20020061543A1
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/873,637
; CURRENT FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 577
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-873-637-2
Alignment Scores:
Pred. No.: 3,41e-150 Length: 577
Score: 2019.50 Matches: 394
Percent Similarity: 78.94% Conservative: 82
Best Local Similarity: 65.34% Mismatches: 96
Query Match: 32.99% Indels: 31
DB: 9 Gaps: 4
US-09-270-437D-6 (1-3412) x US-09-873-637-2 (1-577)
```

```
QY 73 ATGAACAAGCTTTATACATCGGGAACCTGAGCCCGCGTCACCGCCGACGACCTTCGGGAG 132
Db 1 MetAsnLysLeuTyIleGlyAsnLeuAsnGluSerValThrProAlaAspLeuGluLys 20
QY 133 CTCCTTTGGGAGCAGAGAGCTGCCCTGGCGGACAGGTCTCTGTAAGTCCGGCTACGCC 192
Db 21 ValPheAlaGluHisLysIleSerTySerGlyGlnPheLeuValLysSerGlyTyAla 40
QY 193 TTCGTGGACTACCCGACAGAACTGGGCGCATCCGCGCCATCGAGACCTCTCGGGTAAA 252
Db 41 PheValAspCysProAspGluHisTrpAlaMetLysAlaIleGluThrPheSerGlyLys 60
QY 253 GTGGAATTCATCGGGAATATCATGGAAGTTGATTCTACTAGTCTCTATAAAGCTTAAGAC 312
Db 61 ValGluLeuGlnGlyLysArgLeuGluMetGluHisSerValProLysLysGlnArgSer 80
QY 313 AGGAAATTCAGATTCGAAACATCCCTCTCCTACCTGAGTGGAGGTGTGGATGACATT 372
Db 81 ArgLysIleGlnIleArgAsnIleProGlnLeuArgTrpGluValLeuAspSerLeu 100
QY 373 TTGGCTCAATATATGGGACAGTGGAGATGTGGAACTCAACAGACACACAGAAACCGCC 432
Db 101 LeuAlaGlnTyGlyThrValGluAsnCysGluGlnValAsnThrGluSerGluThrAla 120
QY 433 GTTGTCAACGTCAATATGCAACAGAGAGAAAGCAAAATATGCCATGGAGAGCTAAGC 492
Db 121 ValValAsnValThrTySerAsnArgGluGlnThrAlaIleMetLysLeuAsn 140
QY 493 GGGCATCAGTTTGAAGACTACTCCTTCAAGATTTCCTACATCCCGATGAAGAGGTGAGC 552
```

Db 141 GlyHisGlnLeuGluAsnHisAlaLeuLysValSerTyrIleProAspGluGlnIleThr 160
QY 553 TCCCTTCCGCCCCCTCAGCGGCCAGCGTGGGGACCACTCTCCGGGAGAGGCCAC 612
Db 161 -----GlnGlyProGluAsnGlyArgGlyGlyPheGlySerArgGlyGlnProArg 178
QY 613 -----GCCCTCGGGGCACTCTCCAGCCAGCCAGATGATTCCCGCTG 657
Db 179 GlnGlySerProValAlaAlaGlyAlaProAlaLysGlnGlnProValAspIleProLeu 198
QY 658 CGGATCTGGTCCCAACCCAGTTGTGTGTCATCATCGGAAGAGGCGCTTGACCAATA 717
Db 199 ArgLeuLeuValProThrGlnTyrValGlyAlaIleIleGlyLysGluGlyAlaThrIle 218
QY 718 AAGACATCACTAAGCAGACCAGTCCCGGTAGATATCCATAGAAAAGAGAACTCTGA 777
Db 219 ArgAsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGly 238
QY 778 GCTGAGAGAAAGCCTGTCAACATCCATGCCACCCAGAGGGGACTTCTGAAGCATGCCGC 837
Db 239 AlaAlaGluLysAlaIleSerValHisSerThrProGluGlyCysSerSerAlaCysLys 258
QY 838 ATGATCTTGAATCATCAGAAAGAGCAGATGACCAACTAGCCGAGAGAGATTCT 897
Db 259 MetIleLeuGluIleMetHisLysGluAlaLysAspThrLysThrAlaAspGluValPro 278
QY 898 CTGAAATCTGGCACACAATGGCTGTGTGAAGACTGATTGGAAAAGAGCGAAT 957
Db 279 LeuLysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsn 298
QY 958 TTGAAGAAATTTGAACATCAACAGGACCAAGATAACAATCTCATCTTTCAGGATTG 1017
Db 299 LeuLysLysValGluGluAspThrGluThrLysIleThrIleSerSerLeuGlnAspLeu 318
QY 1018 AGCATATACACCCGAAAGAACCATCACTGTGAAGGCACAGTTGAGCCCTGTGCCAGT 1077
Db 319 ThrLeuTyrAsnProGluArgThrIleThrValLysGlyAlaIleGluAsnCysCysArg 338
QY 1078 GCTGAGATAGATATATCAAGACTGCTGAGGCTTTGAAATGATATGCTGCTGTT 1137
Db 339 AlaGluGlnGluIleMetLysLysValArgGluAlaTyrGluAsnAspValAlaAlaMet 358
QY 1138 AACCAACAAAGCAATCTCATCCAGGTTGAACCTCAGCGCATTTGGCATCTTTCAACA 1197
Db 359 SerLeuGlnSerHisLeuIleProGlyLeuAsnLeuAlaValAlaGlyLeuPheProAla 378
QY 1198 GGACTGTCCTGCTATCTCCACAGACAGGCCCCCGGAGCTCCCGCTGCCCCCTAC 1257
Db 379 SerSerSerAlaValProProPro-----ProSerSerValThrGlyAlaAlaProTyr 396
QY 1258 CACCCCTTCACTACCCACTCCGGATCTTCTCCAGCCTGTACCCCGCATCACAGTTTGGC 1317
Db 397 SerSerPheMet----- 400
QY 1318 CCGTTCGCGCATCATCTTATCCAGAGCAGAGATTGTGAATCTCTTTCATCCCAACC 1377
Db 401 -----GlnAlaProGluGlnGluMetValGlnValPheIleProAla 414
QY 1378 CAGCTGTGGCGGCATCATCGGAAGAGGGGCACATCAACAGCTGGCGAGATTC 1437
Db 415 GlnAlaValGlyAlaIleIleGlyLysGlyGlnHisIleLysGlnLeuSerArgPhe 434
QY 1438 GCGGAGCTCTATCAGATTGCCCTCGGAGAGCCAGCAGCTCAGGAAAGATGGTC 1497
Db 435 AlaSerAlaSerIleLysIleAlaProProGluThrProAspSerLysValArgMetVal 454
QY 1498 ATCATCACCGGGCCACCGGAAGCCAGTTCAAGGCCAGAGGACGATCTTTGGGAAACTG 1557
Db 455 ValIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysLeu 474
QY 1558 AAAGAGAAACTTCTTTAAACCCCAAGAGAGTGAAGCTGGAAGCGCATATCAGAGTG 1617
Db 475 LysGluGluAsnPhePheGlyProLysGluGluValLysLeuGluThrHisIleArgVal 494

QY 1618 CCCTCTTCCACAGCTGGCCGGGTGATTGGCAAAAGTGGCAAGCCGTGAACGACTGCAG 1677
Db 495 ProAlaSerAlaAlaGlyArgValIleGlyLysGlyGlyLysThrValAsnGluLeuGln 514
QY 1678 AACTTAACCACTGACAGAGTCACTCGCTGCTGACCAAAACCCAGATGAAAATGAGNA 1737
Db 515 AsnLeuThrAlaAlaGluValValProArgAspGlnThrProAspGluAsnAspGln 534
QY 1738 GTGATCGTCAGAAATATCGGGCAGCTTCTTGTACCCAGACTGCACAGCGCAAGATCAGG 1797
Db 535 ValIleValLysIleIleGlyHisPheTyrAlaSerGlnMetAlaGlnArgLysIleArg 554
QY 1798 GAAATTGTACACAGGTGAACAGCAGAGCAGAGAAATACCTCAGGGAGTGCCTCACAG 1857
Db 555 AspIleLeuAlaGlnValLysGlnHisGlnLysGlyGlnSerAsnLeuAlaGlnAla 574
QY 1858 GCGACGAC 1866
Db 575 ArgArgLys 577
RESULT 8
US-10-313-986-486
; Sequence 486, Application US/10313986
; Publication No. US20030236209A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: McNabb, Andria
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tongcong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C19
; CURRENT APPLICATION NUMBER: US/10/313,986
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 486
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-313-986-486
Alignment Scores:
Pred. No.: 3,64e-146 Length: 589
Score: 1968.50 Matches: 395
Percent Similarity: 77.69% Conservative: 82
Best Local Similarity: 64.33% Mismatches: 106
Query Match: 32.16% Indels: 31
DB: 15 Gaps: 10
US-09-270-437D-6 (1-3412) x US-10-313-986-486 (1-589)
QY 73 ATGACAGCTTTACATCGGGAACTGAGCCCCCGCTACCCGCGCAGCTCCGGCAG 132
Db 1 MetAsnLysLeuTyrIleGlyAsnLeuSerGluAlaProSerAspLeuGluSer 20
QY 133 CTCTTTGGGACAGGAAGCTGCCCTGGCGGACAGCTCCTGCTGAAGTCCGGCTAGCC 192
Db 21 IlePheLysAspAlaLysIleProValSerGlyProPheLeuValLysThrGlyTyrAla 40
QY 193 TTCGTGGACTACCCCGACACCACTGGGCCCATCGGCCCATCGAGACCTCTCGGGTAAA 252
Db 41 PheValAspCysProAspGluSerTrpAlaLeuLysAlaIleGluAlaLeuSerGlyLys 60
QY 253 GTGGAATTCATGGGAAAATCATGGAAGTTGATTCTCAGTCTCTAAAAAGCTAAGAGC 312
Db 61 IleGluLeuHisGlyLysProIleGluValGluHisSerValProLysArgGlnArgIle 80
QY 313 AGGAATTCAGATTCTGAAACATCCCTCCTCAGCTGAGTGGGAGGTGTGGATGACTT 372
Db 81 ArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspSerLeu 100

QY	373	TTGGCTCAATATATGGGACAGTGGAGAACTGTGGAAACAAGTCAACACACAGACACAGAAACCGCC	432
DB	101	LeuValGlnTyrGlyValValGluSerCysGluInValAsnThrAspSerGluThrAla	120
QY	433	GTGTGTCACGTCACATATGCAACAAGAGAGAAGCAAAATAGCCATGGAGAGCTAAGC	492
DB	121	ValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn	140
QY	493	GGGCATCAGTTTCAGAACTACTCCTTCAAGATTCTCATCCCGGATCAAGAGGTGAGC	552
DB	141	GlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAspGluThrAlaAla	160
QY	553	TCGCTTCGCCCTCAG-----CGAGCCCCAGCGT-----GGGAGCCACTTCTCCCGG	600
DB	161	GlnGlnAsnProLeuGlnGlnProArgGlyArgGlyLeuGlyGlnArgGlySerSer	180
QY	601	GAGCAGGCCACGCCCTGGGGCACTTCTCAGGCCACAGAGATTGATTCGCGTGGCGG	660
DB	181	ArgGlnGly--SerProGlySerValSerLysGlnLysProCysAspLeuProLeuArg	199
QY	661	ATCTGTGTCGCCACCCAGTTGTGTGGTGCCTCATCGGAAAGGAGGCGCTTGACCAATAAG	720
DB	200	LeuLeuValProThrGlnPheValGlyAlaIleIleGlyLysGluGlyAlaThrIleArg	219
QY	721	AACATCACTAAGCAGACCCAGTCCGGGTAGATATCCATGAAAGAGAACTCTGAGCT	780
DB	220	AsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGlyAla	239
QY	781	GCAGAGAAGCTGTACCATCCATCGCCACCCAGAGGGAGTCTCTGAAGCATGCCCATG	840
DB	240	AlaGluLysSerIleThrIleLeuSerThrProGluGlyThrSerAlaLacLysLysSer	259
QY	841	ATTCTTGAATCATGCAGAAAGAGCGCAGATGACAGCAACAACTAGCCGAGAGATTCCTCTG	900
DB	260	IleLeuGluIleMetHisLysGluAlaGlnAspIleLysPheThrGluIleProLeu	279
QY	901	AAATCTTGGCACAATGGCTGTGGTGGAGACTGATTCGGAAAGAGCGACAAATTTG	960
DB	280	LysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsnLeu	299
QY	961	AAGAAATTTGAACATGAAACAGGACCAAGATACCAATCTCATCTTTGCAGGATTTTGGC	1020
DB	300	LysLysIleGluGlnAspThrAspThrLysIleThrIleSerProLeuGlnGluLeuThr	319
QY	1021	ATATACAACCGGAAAGAACCATCAGCTGTGAAGGGCACAGCTTGAGGCGCTGTGCCAGTGT	1080
DB	320	LeuTyrAsnProGluArgThrIleThrValLysGlyAsnValGluThrCysAlaLysAla	339
QY	1081	GAGATAGAGATTATGAAGAGCTCGGTGAGCGCTTTGAAAATCATATGCTGGCTGTTAAC	1140
DB	340	GluGluGluIleMetLysLysIleArgGluSerTyrGluAsnAspIleAlaSerMetAsn	359
QY	1141	CAACAAGCAATCTGATCCAGGGTTGAACTCAGCGCACTTGGCATCTTTTCAACAGGA	1200
DB	360	LeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeuGlyLeuPhe-----	376
QY	1201	CTGTCCGTGTATCTCCACACAGAGGGCCCGGGAGCTCCCCCGCTGCCCTACCCAC	1260
DB	377	-----ProProThrSer-----GlyMetProProThr-----	386
QY	1261	CCCTTCACTACCACTCGGNATCTTCCAGCTGTACCCCATCACAGTTTGGCCCG	1320
DB	387	-----SerGlyProProSerAlaMetThrPro-----Pro	396
QY	1321	TTCCCGCATCATCACTCTTATCCAGAGCAGAGATTCTGAATCTTCTTCATCCCAACCCAG	1380
DB	397	TyrProGlnPheGluGln--SerGluThrGluThrValHisLeuPheIleProAlaLeu	415
QY	1381	GCTGTGGGGCGCATCATCGGGAAGAGGGGGCACACATCAAAACAGCTGGCGAGATTGCC	1440
DB	416	SerValGlyAlaIleIleGlyLysGlnHisIleLysGlnLeuSerArgPheAla	435

Qy	1441	GGAGCCTCTATCAAGATTGCCCCCTCGGGAAAGGCCCAAGACGTCAAGCAAGAGGATGGTCATC	1500
Db	436	GlyAlaSerIleIysIleAlaProAlaGluAlaProAspAlaIysValArgMetValIle	455
Qy	1501	ATCACGGGCGCACCGGAAGCCCAAGTTTCAAGGCCCGAGGCGGATCTTTGGGAAACTGAAA	1560
Db	456	IleThrGlyProGluAlaGlnPheLysAlaGlnGlyArgIleIyrGlyIysIleLys	475
Qy	1561	GAGGAAAACTTCTTTAAACCCCAAGAAAGTGAAGTCGAAGCGCATATCAGAGTGCCC	1620
Db	476	GluGluAsnPheValSerProLysGluGluValLysLeuGluAlaHisIleArgValPro	495
Qy	1621	TCTTCCACACTCGCGGGTGATTGGCAAGGTGGCAAGCCGTGAACGAACTGCAGAAC	1680
Db	496	SerPheAlaAlaGlyArgValIleGlyLysGlyGlyLysThrValAsnGluLeuGlnAsn	515
Qy	1681	TTAAACAGTCGACAGTAAGTCATCGTCCTCGGTGACCAACAGCCACGATGAAAATCAGGAAGTG	1740
Db	516	LeuSerSerAlaGluValValProArgAspGlnThrProAspGluAsnAspGlnVal	535
Qy	1741	ATCGTCGAGAATATCGGGCACTTCTTTGTAGCCACACTGCACAGCGCAAGATCAGGGAA	1800
Db	536	ValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGlnArgIysIleGlnGlu	555
Qy	1801	ATTGTACACAGGTGAAGCAG---CAGGACGACAAATACCCCTCAG---GGAGTCGCCTCA	1854
Db	556	IleLeuThrGlnValLysGlnHisGlnGlnGlnLysAlaLeuGlnSerGlyProProGln	575
Qy	1855	CAGCGCAGCAAGTCAGGCTCCCAAGCAGCCACCAACAAAC	1896
Db	576	SerArgArgLysHisHisHisHisHisHisHisHisHisHis	589
RESULT 9			
US-09-735-705-348			
; Sequence 348, Application US/09735705			
; Patent No. US20020052329A1			
; GENERAL INFORMATION:			
; APPLICANT: Wang, Tongtong			
; APPLICANT: Fan, Liqun			
; APPLICANT: Kalos, Michael D.			
; APPLICANT: Bangor, Chaitanya S.			
; APPLICANT: Hosken, Nancy			
; APPLICANT: Fanger, Gary R.			
; APPLICANT: Li, Samuel X.			
; APPLICANT: Wang, Aijun			
; APPLICANT: Skeiky, Yasir A.W.			
; APPLICANT: Henderson, Robert A.			
; APPLICANT: McNeill, Patricia D.			
; APPLICANT: Fanger, Neil			
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY			
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER			
; FILE REFERENCE: 210121.455C14			
; CURRENT APPLICATION NUMBER: US/09/735,705			
; CURRENT FILING DATE: 2000-12-12			
; NUMBER OF SEQ ID NOS: 419			
; SOFTWARE: FastSeq for Windows Version 3.0			
; SEQ ID NO 348			
; LENGTH: 579			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-09-735-705-348			
Alignment Scores:			
Pred. No.:		2,67e-145	Length: 579
Score:		1957.50	Matches: 387
Percent Similarity:		79.19%	Conservative: 81
Best Local Similarity:		65.48%	Mismatches: 94
Query Match:		31.98%	Indels: 29
DB:		9	Gaps: 8
US-09-270-437D-6 (1-3412) x US-09-735-705-348 (1-579)			


```

Db      1 MetAsnLysLeuTyrlleGlyAsnLeuSerGluAsnAlaIaProSerAspLeuGluSer 20
QY      133 CTCITTTGGGACAGGAAGCTGCCCTTGGGGACAGGTCCTGCTGAAATGTCGGGTACGCC 192
Db      21 IlePheLysAspAlaLysIleProValSerGlyProPheLeuValLysThrGlyTyrAla 40
QY      193 TCGTGGGACTACCCGACAGAACCTGGCCCATCGCGCATCGAGACCCTCTCGGGTAAA 252
Db      41 PheValAspCysProAspGluSerIrpAlaLeuLysAlaIleGluAlaLeuSerGlyLys 60
QY      253 GTGGAATTCATGGGAAAATCATGGAAGTTGATTCTACTCAGTCTCTTAAAAAAGCTAAGGAGC 312
Db      61 IleGluLeuHisGlyLysProIleGluValGluHisSerValProLysArgGlnArgIle 80
QY      313 AGGAAAATTCAGATTCGAAACATCCCTCTCTACCTCGAGTGAGGAGGTGTGGATGACTT 372
Db      81 ArgLysLeuGlnIleArgAsnIleProHisLeuGlnTrpGluValLeuAspSerLeu 100
QY      373 TTGGCTCAATATGGGACAGTGGAGAAATGTGGAACAAGTCAACACAGACACAGAACCGCC 432
Db      101 LeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAspSerGluThrAla 120
QY      433 GTTGTCACAGTCACATATCAACAGAGAGAAGCAAAAATAGCCATCGAGAAGCTAAGC 492
Db      121 ValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn 140
QY      493 GGCCATCAGTTTGAGAACTACTCCTTCAAGATTTCCTACATCCCGGATGAAGAGGTGAGC 552
Db      141 GlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAspGluThrAlaAla 160
QY      553 TCCCTCTCGCCCTCCAG- ----CGAGCCAGAGT- ----GGGGACCATCTCTCCCGG 600
Db      161 GlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlyGlnArgGlySerSer 180
QY      601 GAGCAAGCCACGCCCTTGGGGCACTTCTCAGGCCACAGATGTGATTTCCCGTGGCG 660
Db      181 ArgGlnGly- --SerProGlySerValSerLysGlnLysProCysAspLeuProLeuArg 199
QY      661 ATCTGTGTCGCCACCCAGTTTGTGGTCCCATCATCGAAAGAGGGGCTTGACCATAAAG 720
Db      200 LeuLeuValProThrGlnPheValGlyAlaIleIleGlyLysGluGlyAlaThrIleArg 219
QY      721 AACATCACTAAGCAGACCCAGTCCCGGTGATATCCATAGAAAAGAGAACTCTCGAGCT 780
Db      220 AsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGlyAla 239
QY      781 GCAGAGAAGCTGTACCATCCATCCACCCAGAGGGGACTTCTGAAGCATGCGCGATG 840
Db      240 AlaGluLysSerIleThrIleLeuSerThrProGluGlyThrSerAlaAlaCysLysSer 259
QY      841 ATTCTTGAATCATGCAGAAAGAGGCAGATGAGACCAAACTAGCGGAGAGATTCCTCTG 900
Db      260 IleLeuGluIleMetHisLysGluAlaGlnAspIleLysPheThrGluGluIleProLeu 279
QY      901 AAAATCTTTGGCACACAATGGCTTGGTGGAGACTGATTGCAAAAGAGGAGGAGAAATTG 960
Db      280 LysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsnLeu 299
QY      961 AAGAAAATTGAACATGAACAGGACCAAGATACAAATCTCATCTTTGCAGGATTTGAGC 1020
Db      300 LysLysIleGluGlnAspThrAspThrLysIleThrIleSerProLeuGlnGluLeuThr 319
QY      1021 ATATACACCCGGAAAGAACCATCATCTGTGAAGGCACAGTTGAGGGCTGTGCCAGTGCT 1080
Db      320 LeuTyrAsnProGluArgThrIleThrValLysGlyAsnValGluThrCysAlaLysAla 339
QY      1081 GAGATAGAGATTATGAAGAAGCTCGTGAGGCTTTTGAATAATGATATGCTGGCTGTAAAC 1140
Db      340 GluGluGluIleMetLysLysIleArgGluSerTyrGluAsnAspIleAlaSerMetAsn 359
QY      1141 CAAACAGCCAAATCTGATCCCGAGGTGAACCTCAGCGCACTTGGCATCTTTTCAACAGGA 1200

```

```

Db      360 LeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeuGlyLeuPhe- ---- 376
QY      1201 CTGTCCTGTGTATCTTCACACAGCAGGGCCCGCGAGCTCCCGCCGCTGCCCTTACCAC 1260
Db      377 - ----ProThrSer- ----GlyMetProProProThr- ---- 386
QY      1261 CCCITCACTACCCACTCGGATCTTCTCAGCCTGTACCCCATCACACAGTTTGGCCCG 1320
Db      387 - ----SerGlyProProSerAlaMetThrPro- ----Pro 396
QY      1321 TTCCCGCATCATCTCTTATCCAGAGCAGAGATTGTGAATCTCTTATCCCAACCCAG 1380
Db      397 TyrProGlnPheGluGln- --SerGluThrGlnThrValHisLeuPheIleProAlaLeu 415
QY      1381 GCTGTGGCGCCATCATCGGAAGAGGGGCACACATCAACAGACTGGCGGAGATTTCGCC 1440
Db      416 SerValGlyAlaIleIleGlyLysGlnGlyGlnHisIleLysGlnLeuSerArgPheAla 435
QY      1441 GGAGCCTCTATCAAGATTGCCCTCGCGAAGGCCAGACGTGAGGAAAGGATGGTCTATC 1500
Db      436 GlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLysValArgMetValIle 455
QY      1501 ATCAACCGGGCCACCGAAGCCAGTTCGAAGCCCGAGGAGCGATCTTTGGGAAACTGAAA 1560
Db      456 IleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysIleLys 475
QY      1561 GAGAAAACCTTCTTTAAACCCCAAGAGAGTGAAGCTGGAAGCGCATATCAGAGTGGCC 1620
Db      476 GluGluAsnPheValSerProLysGluValLysLeuGluAlaHisIleArgValPro 495
QY      1621 TCTTCCACAGCTGCGCGGTGATTGGCAAGGTGGCAGACCGTGAACGAACTGAGAAC 1680
Db      496 SerPheAlaAlaGlyArgValIleGlyLeGlyLysThrValAsnGluLeuGlnAsn 515
QY      1681 TTAACCAAGTCAGAAAGTCATCGTCTCGTACCAAAAGCCAGATGAAATAGAGGAGTG 1740
Db      516 LeuSerSerAlaGluValValProArgAspGlnThrProAspGluAsnAspGlnVal 535
QY      1741 ATCTCAGAAATTATCGGCGCACTTCTTTGCTAGCCAGACTGCACAGCGCAGATCAGGAA 1800
Db      536 ValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGlnArgLysIleGlnGlu 555
QY      1801 ATTCTACAACAGGTGAAGCAGCAGGACAGAGAA 1833
Db      556 IleLeuThrGlnValLysGlnHisGlnGlnGln 566

RESULT 10
US-09-850-716A-348
; Sequence 348. Application US/09850716A
; Patent No. US20020115139A1
; GENERAL INFORMATION:
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Retter, Marc W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C15
; CURRENT APPLICATION NUMBER: US/09/850,716A
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 348
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-850-716A-348

Alignment Scores:
Pred. No.: 2,67e-145 Length: 579
Score: 1957.50 Matches: 387
Percent Similarity: 79.19% Conservative: 81
Best Local Similarity: 65.48% Mismatches: 94
Query Match: 31.98% Indels: 29

```

DB: 9 Gaps: 8
US-09-270-437D-6 (1-3412) x US-09-850-716A-348 (1-579)
QY 73 ATGAACAAGCTTTACATCGGGAACCTGAGCCCGCGCTCACCGCCGACGACCTCCGGCAG 132
DB 1 MetAsnLysLeuTyrIleGlyAsnLeuSerGluAsnAlaAlaProSerAspLeuGluSer 20
QY 133 CTCCTTTGGGGACAGAAAGCTGCCCTCGGGGACAGAGTCTCTGCTGAAGTCCGGCTACGCC 192
DB 21 IlePheLysAspAlaLysIleProValSerGlyProPheLeuValLysThrGlyTyrAla 40
QY 193 TTCGTGGACTACCCGACAGAACTGGGCCATCCGCGCCATCGAGACCTCTCGGCTAA 252
DB 41 PheValAspCysProAspGluSerIrpAlaLeuLysAlaIleGluAlaLeuSerGlyLys 60
QY 253 GTGGAAATTCATCGGGAATCATGGAAGTTGATTAATCTCACTCTCTAAAAGTAAGGAGC 312
DB 61 IleGluLeuHisGlyLysProIleGluValGluHisSerValProLysArgGlnArgIle 80
QY 313 AGGAAATTCAGATTGCAAACTCCCTCCTCACTGCGAGTGGAGGTGTGGATGACTT 372
DB 81 ArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspSerLeu 100
QY 373 TTGGCTCAATATGGGACAGTGGAGATGTGGAACAAGTCAACACACACAGAAACCGCC 432
DB 101 LeuValGlnTyr-GlyValValGluSerCysGluGlnValAsnThrAspSerGluThrAla 120
QY 433 GTTGTCAACGTCATATGCAACAAGAGAAAGCAAAATAGCCATGGAGAGCTTAAGC 492
DB 121 ValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn 140
QY 493 GGGCATCATGTTGAGAACTACTCCTCAAGATTCTCATATCCGATCGAGGTGAGC 552
DB 141 GlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAspGluThrAlaAla 160
QY 553 TCCCTTCCTCCCTCCTCAG-----CGAGCCACGCT-----GGGACCACTCTCCCGG 600
DB 161 GlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlyGlnArgGlySerSer 180
QY 601 GAGCAAGGCCACGCCCTCGGGGCACTTCTCAGGCCAGACAGATTGATTTCCGCTGGCG 660
DB 181 ArgGlnGly---SerProGlySerValSerLysGlnLysProCysAspLeuProLeuArg 199
QY 661 ATCTCGTCCACCCAGTTCTTGTGGCCATCATCGAAAGAGAGGCTTGACCATTAAG 720
DB 200 LeuLeuValProThrGlnPheValGlyAlaIleIleGlyLysGluGlyAlaThrIleArg 219
QY 721 AACATCACTAAGCAGACCCCGGCTAGATATCCATAGAAAAGAGAACTCTGGAGCT 780
DB 220 AsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGlyAla 239
QY 781 GCAGAGAAGCTGTACCATTCATCCATGCCACCCAGAGGGACTTCTGAAGATGCCCATG 840
DB 240 AlaGluLysSerIleThrIleLeuSerThrProGluGlyThrSerAlaAlaCysLysSer 259
QY 841 ATTCTTGAATCATGCAGAAAGAGGCGAGATGAGACCAAACTAGCCCAAGAGATTCTCTG 900
DB 260 IleLeuGluIleMethIleLysGluAlaGlnAspIleLysPheThrGluGluIleProLeu 279
QY 901 AAAATCTTGGCCACAAATGGCTTGGTTGGAAGACTGATTGGAAGAGAGCGAGAAATTTG 960
DB 280 LysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsnLeu 299
QY 961 AGAAAATGAACATGAACAGGACCAAGATTAACAATCTCATCTCTTTCAGAGTTTGAGC 1020
DB 300 LysLysIleGluGlnAspThrAspThrLysIleThrIleSerProLeuGlnGluLeuThr 319
QY 1021 ATATACACCCGGAAGAACCATCACTGTGCAAGGGCACAGTTCAGGCTGTGCGACTGCT 1080
DB 320 LeuTyrAsnProGluArgThrIleThrValLysGlyAsnValGluThrCysAlaLysAla 339
QY 1081 GAGATAGAGATTATGAAGAAGCTGCTGAGGCCCTTTGAAAATGATATGCTGGCTGTTAAC 1140

DB 340 GluGluGluIleMetLysLysIleArgGluSerTyrGluAsnAspIleAlaSerMetAsn 359
QY 1141 CAACAAGCAATCTCATCCAGGGTTGAACCTCAGCGACCTTGGCATCTTTTCAACAGA 1200
DB 360 LeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeuGlyLysPhe----- 376
QY 1201 CTGTCTGTCTATCTCCACACAGAGGGCCCCCGCGAGCTCCCCCGCTGCCCTTACCAC 1260
DB 377 -----ProProThrSer-----GlyMetProProProThr----- 386
QY 1261 CCCTTCACTACCCACTCCGGATACTTCTCCAGCTTACCCCATCACAGTTTGGCCCG 1320
DB 387 -----SerGlyProProSerAlaMetThrPro-----Pro 396
QY 1321 TTCCTGCATCATCTCTTATCCAGAGCAGAGATTGAAATCTCTCATCCCAACCCAG 1380
DB 397 TyrProGlnPheGluGln---SerGluThrGluThrValHisLeuPheIleProAlaLeu 415
QY 1381 GCTGTGGCGCCCATCATCGGAAGAGGGGACACATCAACAGCTGGCGAGATTCCGC 1440
DB 416 SerValGlyAlaIleIleGlyLysGlnGlnHisIleLysGlnLeuSerArgPheAla 435
QY 1441 CGAGCTCTTATCAAGATTGCCCTCGGAAGCCCGACAGCTCAGCGAAAAGGATGTCATC 1500
DB 436 GlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLysValArgMetValIle 455
QY 1501 ATCAGCGGCCACCGAGCCCGAGTTCAGGCGCCAGGACCGATCTTTGGGAAACTGAAA 1560
DB 456 IleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysIleLys 475
QY 1561 GAGGAAAACCTCTTTAAACCCCAAGAAAGTGAAGCTGGAAGCGCATATCAGAGTGC 1620
DB 476 GluGluAsnPheValSerProLysGluGluValLysLeuGluAlaHisIleArgValPro 495
QY 1621 TCTTCCACAGCTGGCGGTGATTGGCAAGGTGCAGACCGTGAACCAACTGCGAAGC 1680
DB 496 SerPheAlaAlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsn 515
QY 1681 TTAAACAGTGCAGAGTCATCTGCTCTGTCGACCAACCCAGATCAAAATGAGGAAGT 1740
DB 516 LeuSerSerAlaGluValValProArgAspGlnThrProAspGluAsnAspGlnVal 535
QY 1741 ATCTGTCAAGATTATCGGCACTCTTGTAGCCAGACTGCAGCGCAAGATCAGGGAA 1800
DB 536 ValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGlnArgLysIleGlnGlu 555
QY 1801 ATTGTACACAGCTGAAGCAGCAGGAGCAGAAA 1833
DB 556 IleLeuThrGlnValLysGlnHisGlnGlnGln 566
RESULT 11
US-09-897-778-348
; Sequence 348, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Pecham, David W.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C16
; CURRENT APPLICATION NUMBER: US/09/897,778
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 348

```
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-897-778-348

Alignment Scores:
  Pred. No.: 2,67e-145      Length: 579
  Score: 1957,50           Matches: 387
  Percent Similarity: 79.19%      Conservative: 81
  Best Local Similarity: 65.48%    Mismatches: 94
  Query Match: 31.98%            Indels: 29
  DB: 9                        Gaps: 8

US-09-270-437D-6 (1-3412) x US-09-897-778-348 (1-579)

QY 73 ATGAACAAGCTTTACATCGGAACCTGAGCCCGCGCTACCGCGCAGACCTCCGGCAG 132
Db 1 MetAsnLysLeuTyrIleGlyAsnLeuSerGluAsnAlaAlaProSerAspLeuGluSer 20
QY 133 CTCCTTTGGGACAGGAAGCTCCCTCGCGGGACAGGTCCTGCTGAAGTCGGCTACGCC 192
Db 21 IlePheLysAspAlaLysIleProValSerGlyProPheLeuValLysThrGlyTyrAla 40
QY 193 TCTGTGGACTACCCCGACAGAACTGGGCCATCCGCGCATCGACACCTCTCGGGTAAA 252
Db 41 PheValAspCysProAspGluSerTrpAlaLeuLysAlaIleGluAlaLeuSerGlyLys 60
QY 253 GTGGAAATTCATGGGAAATCATGGAATGATGATCTACTCAGTCTCTAAAGACTAAGGAGC 312
Db 61 IleGluLeuHisGlyLysProIleGluValGluHisSerValProLysArgGlnArgIle 80
QY 313 AGGAAATTCAGATTCGAACATCCCTCCTCACCTGCACTGGAGTGGAGTCTTGATGATTT 372
Db 81 ArgLysLeuGlnIleArgAsnIleProProHisLeuGlnIleProGluValLeuAspSerLeu 100
QY 373 TTGGCTCAATATGGGACAGTGGAGAAATGTGAAACAAGTCAACACAGACACAGAAACCGCC 432
Db 101 LeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAspSerGluThrAla 120
QY 433 GTTGTCAAGTCACATATGCAACAGAGAGAGAGCAAAATATAGGCATGGAGAACTAAGC 492
Db 121 ValValAsnValThrTyrSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn 140
QY 493 GGGCATCAGTTTGAGAACTACTCCTTCAGATTTCCTACATCCCGCATGAGAGAGTGGAGC 552
Db 141 GlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAspGluThrAlaAla 160
QY 553 TCCCTCTCGCCCTCCAG-----CGAGCCCGAGCGT-----GGGGACCACTCTCCCGG 600
Db 161 GlnGlnAsnProLeuGlnProArgGlyArgArgGlyLeuGlyGlnArgGlySerSer 180
QY 601 GAGCAAGCCACGCCCTCGGGGCACCTTCTCAGCCGACAGATGATGATTTCCCGCTCGG 660
Db 181 ArgGlnGly----SerProGlySerValSerLysGlnLysProCysAspLeuProLeuArg 199
QY 661 ATCTGTGTCGCCACCCAGCTTTGGTGGCCATCATCGGAAGAGGGCGCTTACCATTAAG 720
Db 200 LeuLeuValProThrGlnPheValGlyAlaIleIleGlyLysGluGlyValThrIleArg 219
QY 721 AACATCACTAAGCAGACCCAGCTCCCGGTAGATATCCATAGAAAAGAGAACTCTGGAGCT 780
Db 220 AsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGlyAla 239
QY 781 GCAGAGAAGCTGTCCACATCATGCCACCCAGAGGGGACTTCTGAAGCATGCCGCATG 840
Db 240 AlaGluLysSerIleThrIleLeuSerThrProGluGlyThrSerAlaAlaCysLysSer 259
QY 841 ATTCTTGAATCATGCAGAAAGAGGCAGATGAGACCAACTAGCCGAGAGATTCCTCTG 900
Db 260 IleLeuGluIleMetHisLysGluAlaGlnAspIleLysPheThrGluGluIleProLeu 279
QY 901 AAAATCTTTGGCACCAATATGGCTTGGTGGAAAGACTGATTTGAAAAGAGAGGAGAAATTG 960
Db 960
```

```
Db 280 LysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsnLeu 299
QY 961 AAGAAAATTGAACATGAACAGGACCAAGGATAACAATCTCTATCTTTTCAGGATTTGAGC 1020
Db 300 LysLysIleGluGlnAspThrAspThrLysIleThrIleSerProLeuGlnGluLeuThr 319
QY 1021 ATATACAACCCGGAAGAAAGAACCATCACTGTGAAGGGCACAGTTGAGCCCTGTGCCAGTCT 1080
Db 320 LeuTyrAsnProGluArgThrIleThrValLysGlyAsnValGluThrCysAlaLysAla 339
QY 1081 GAGATAGAGATTATGAAGAAGCTCGTGAGCCCTTTGAAAATGATATATCTGCTGTGTTAAC 1140
Db 340 GluGluGluIleMetLysLysIleArgGluSerTyrGluAsnAspIleAlaSerMetAsn 359
QY 1141 CAACAAGCAATCTCATCCAGGGTTGAACCTCAGCGCACTTGGCACTTTTTCACAGGA 1200
Db 360 LeuGlnAlaHisLeuIleProGlyLeuAsnLeuAlaLeuGlyLeuPhe----- 376
QY 1201 CTGTCCGTGTATCTCCACAGAGAGGGCCCGCGGAGCTCCCGCGTGGCCCCCTACAC 1260
Db 377 -----ProProThrSer-----GlyMetProProProThr----- 386
QY 1261 CCTTCACTACCCACTCCGATCTCTCCAGCCTGTACCCCATCACCAGTTTGGCCCG 1320
Db 387 -----SerGlyProProSerAlaMetThrPro-----Pro 396
QY 1321 TTCCCGCATCATCTCTTATCCAGAGAGGATTTGAATCTTCTATCTTATCCCAACCCAG 1380
Db 397 TyrProGlnPheGluGln---SerGluThrGluThrValHisLeuPheIleProAlaLeu 415
QY 1381 GCTGTGGCGGCATCATCTCGGGAAGAGAGGGGCACACATCAACAGCTGGCAGATTCGCC 1440
Db 416 SerValGlyAlaIleIleGlyLysGlnGlyHisIleLysGlnLeuSerArgPheAla 435
QY 1441 GGAGCCTCTATCAAGATTGCCCTCGGAGAGGGCCAGCGTCACGAAAGGATGTCATC 1500
Db 436 GlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLysValArgMetValIle 455
QY 1501 ATCACCGGCCACCGGAGAGCCAGTTTCAAGSCCCAGGACGGATCTTTGGGAAACTGAAA 1560
Db 456 IleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrGlyLysLys 475
QY 1561 GAGGAAAACCTCTTTAACCCCAAGAGAGAGTGAAGCTGGAAGCGCATATCAGAGTGC 1620
Db 476 GluGluAsnPheValSerProLysGluGluValLysLeuGluAlaHisIleArgValPro 495
QY 1621 TCTTCCACAGCTGCCCGGTGATTGGCAAGGTGGCAAGACCGTGAACGAACTGCAGAAC 1680
Db 496 SerPheAlaAlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsn 515
QY 1681 TTAACCGAGTCGAGAAGTCATCGTCCCTCGTGACCAAAACGCCAGATGAAAATGAGGAAGTG 1740
Db 516 LeuSerSerAlaGluValValProArgAspGlnThrProAspGluAsnAspGlnVal 535
QY 1741 ATCTGTAGAATTATCGGGCACTCTTCTGTAGCCAGACTGCACAGCGCAAGATCAGGAA 1800
Db 536 ValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGlnArgLysIleGlnGlu 555
QY 1801 ATTGTACAACAGGTGAAGCAGCAGCAGCAAGAA 1833
Db 556 IleLeuThrGlnValLysGlnHisGlnGlnGln 566
```

RESULT 12

```
US-09-897-778-446
; Sequence 446, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Mainerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Barrick
; APPLICANT: Watanabe, Yoshihiro
```

```
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C16
; CURRENT APPLICATION NUMBER: US/09/897,778
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 446
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-897-778-446

Alignment Scores:
Pred. No.:      2,67e-145      Length:      579
Score:          1957.50        Matches:    387
Percent Similarity: 79.19%      Conservative: 81
Best Local Similarity: 65.48%    Mismatches: 94
Query Match:      31.98%        Indels:     29
DB:               9            Gaps:       8

US-09-270-437D-6 (1-3412) x US-09-897-778-446 (1-579)
QY 73 ATGAACAAGCTTTACATCGGACCTGAGCCCGCGCTCACCCGCCAGACCTCCGGCAG 132
Db 1 MetAsnLysLeuTyrlleGlyAsnLeuSerGluAsnAlaProSerAspLeuGluSer 20
QY 133 CTCCTTTGGGACAGGAAGTGCCTCGCGGACAGCTCTGCTGAAGTCCCGCTACGCC 192
Db 21 IlePheLysAspAlaLysIleProValSerGlyProPheLeuValIysThrGlyTyAla 40
QY 193 TTCGTGGACTACCCGACACAGAACTGGGCCATCCGCGCATCGAGACCCTCTCGGTAA 252
Db 41 PheValAspCysProAspGluSerTrpAlaLeuLysAlaIleGluAlaLeuSerGlyLys 60
QY 253 GTGGAAATTCATGCGAAATCATGGAAGTTGATTACTCAGTCTCTAAAAGCTAAGGAC 312
Db 61 IleGluLeuHisGlyLysProIleGluValGluHisSerValProLysArgGlnArgIle 80
QY 313 AGGAAATTCAGATTGCAACATCCCTCTCCTCAGTGGAGGTGTGGATGACTT 372
Db 81 ArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspSerLeu 100
QY 373 TTGGCTCATATGGACAGTGGAGAATGTGGAACAGTCAACACAGACACAGAAACCGCC 432
Db 101 LeuValGlnTyrlleGlyValValGluSerCysGluGlnValAsnThrAspSerGluThrAla 120
QY 433 GTTGTCAAGCTCACATATCAACAGAGAAGCAAAATAGCCATGCGAGACTAAGC 492
Db 121 ValValAsnValThrTyrlleSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn 140
QY 493 GGGCATCAGTTTGAAGACTACTCTTCAAGATTCTCATATCCCGATGAAGAGTGAGC 552
Db 141 GlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrlleProAspGluThrAlaAla 160
QY 553 TCCCTTCCGCCCTCAG-----CGAGCCAGCGT-----GGGACCACCTCTCCCGG 600
Db 161 GlnGlnAsnProLeuGlnGlnProArgGlyArgGlyArgGlyLeuGlyGlnArgGlySerSer 180
QY 601 GAGCAAGGCCAGCCCTCGGGGCATCTCTCAGGCCAGCAGAGATTGATTCCCGCTGGG 660
Db 181 ArgGlnGly---SerProGlySerValSerLysGlnLysProCysAspLeuProLeuArg 199
QY 661 ATCTGGTCCCAACCCAGTTTGTGTGCCATCATCGGAAAGAGGGCTTGACCATTAAG 720
Db 200 LeuLeuValProThrGlnPheValGlyAlaIleGlyLysGluGlyAlaThrIleArg 219
QY 721 AACATCACTAAGCAGCCAGTCCCGGTAGATATCCATAGAAAAGAGAACTCTCGAGCT 780
Db 220 AsnIleThrLysGlnThrGlnSerIleAspValHisArgLysGluAsnAlaGlyAla 239
```

```
QY 781 GCAGAGAACCTGTGTACCATCCATCCATGCCACCCCGAGAGGAGCTTCTCAACATCCCGCATG 840
Db 240 AlaGluLysSerIleThrIleLeuSerThrProGluGlyThrSerAlaAlaCysLysSer 259
QY 841 ATTTCTTGAATCATCGAGAAAGAGCAGATGAGACCAACTAGCCGCAAGAGATTCTCTCG 900
Db 260 IleLeuGluIleMetHisLysGluAlaGlnAspIleLysPheThrGluGluLeuProLeu 279
QY 901 AAAATCTTGGCACAAATGGCTTGTGGAAGACTGATTGGAAGAAAGAGCGAGAATTGTG 960
Db 280 LysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsnLeu 299
QY 961 AAGAAATGTGAACATGAACAGGACCAAGATAACAATCTCATCTCTTTCAGAGATTGAGC 1020
Db 300 LysLysIleGluGlnAspThrLysIleThrIleSerProLeuGlnGlnLeuThr 319
QY 1021 ATATACAAACCCGAAAGAACCATCACTGTGAAGGCGACAGTTGAGCGCTGTGCCAGTGT 1080
Db 320 LeuTyrlleAsnProGluArgThrIleThrValLysGlyAsnValGluThrCysAlaLysAla 339
QY 1081 GAGATAGAGATTATGAAGAGCTGCTGAGGCTTTGAAATGATATGCTGCTGTTAATC 1140
Db 340 GluGluGluIleMetLysIleArgGluSerTyrlleGluAsnAspIleAlaSerMetAsn 359
QY 1141 CAACAAGCAATCTCATCCAGGGTTGAACCTCAGCGACTTGGCATCTTTTCAACAGGA 1200
Db 360 LeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeuGlyLysPhe----- 376
QY 1201 CTGTCGCTGTATCTCCACAGCAGCGGCCCGCGGAGCTCCCGCGCTCCCGCTACCCAC 1260
Db 377 -----ProProThrSer-----GlyMetProProProThr----- 386
QY 1261 CCCTTCACTACCACTCCGGATCTTCTCCAGCTGTGTACCCCATCACAGTTTGGCCCG 1320
Db 387 -----SerGlyProProSerAlaMetThrPro-----Pro 396
QY 1321 TTCCCGCATCATCACTTATCCAGCAGGAGATTGTGAATCTTCTTATCCCAACCCAG 1380
Db 397 TyrProGlnPheGluGln---SerGluThrGluThrValHisLeuPheIleProAlaLeu 415
QY 1381 GCTGTGGCGCCCATCATCGGAAAGAGGGGCGACACATCAAAACAGCTGGCGAGATTCCGC 1440
Db 416 SerValGlyAlaIleIleGlyLysGlnGlyGlnHisIleLysGlnLeuSerArgPheAla 435
QY 1441 GGAGCTCTATCAAGATTGCCCTCGGAGGCCAGAGCTCAGCGAAAGGATGCTCATC 1500
Db 436 GlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLysValArgMetValIle 455
QY 1501 ATCAGCGGCGCCACCGAGGCCAGTTCAGAGGCCAGGACGGATCTTTGGGAAACTGAAA 1560
Db 456 IleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrlleGlyLysIleLys 475
QY 1561 GAGGAAACTTCTTTAAACCCCAAGAAAGTGAAGCTGGAAGCGCATATCAGAGTGCC 1620
Db 476 GluGluAsnPheValSerProLysGluGluValLysLeuGluAlaHisIleArgValPro 495
QY 1621 TCTTCCAGCTGGCGGGGTGATTGGCAAGTGGCAGACCGCTGAACCACTGCAGAAC 1680
Db 496 SerPheAlaAlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsn 515
QY 1681 TTAAACAGTGCAGAAAGTTCATCGTGTGTGACCAAAAGCCAGATGAAATGAGGAAGTG 1740
Db 516 LeuSerSerAlaGluValValProArgAspGlnThrProAspGluAsnAspGlnVal 535
QY 1741 ATCTGTCAGAAATTATCGGCGACTTCTTGTGTCACCGAGACTGCAGCGCAAGATCAGGAA 1800
Db 536 ValValLysIleThrGlyHisPheTyrlleAlaCysGlnValAlaGlnArgLysIleGlnGlu 555
QY 1801 ATTTGTACAAACAGGTGAAGCAGCAGCAGCAGAAA 1833
Db 556 IleLeuThrGlnValLysGlnHisGlnGlnGln 566
```

RESULT 13

US-09-897-778-449
; Sequence 449, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darriek
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C16
; CURRENT APPLICATION NUMBER: US/09/897,778
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 449
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-897-778-449

Alignment Scores:

Prod. No.:	2.67e-145	Length:	579
Score:	1957.50	Matches:	387
Percent Similarity:	79.19%	Conservative:	81
Best Local Similarity:	65.48%	Mismatches:	94
Query Match:	31.98%	Indels:	29
DB:	9	Gaps:	8

US-09-270-437D-6 (1-3412) x US-09-897-778-449 (1-579)

QY	73	ATGAACAAGCTTTACATCGGAACCTGAGCCCGCCCGCTCACCGCGAGCCTCCGGCAG	132
DB	1	MetAsnLysLeuTyrlleGlyAsnLeuSerGluAsnAlaProSerAspLeuGluSer	20
QY	133	CTCTTTGGGGACAGGAAGCTGCCCTGGCGGACAGAGTCTGCTGAAGTCCGGGTACGCC	192
DB	21	IlePheLysAspAlaLysIleProValSerGlyProPheLeuValLysThrGlyTyra	40
QY	193	TGCTGGACTACCCGACAGACACTGGGCCATCCGCGCATCGAGACCCCTTCGGGTAAA	252
DB	41	PheValAspCysProAspGluSerTrpAlaLeuLysAlaIleGluAlaLeuSerGlyLys	60
QY	253	GTGGAATTGCATGGGAAATCATGAAGTTGATTACTCAGTCTCTAAAAGCTAAGGAGC	312
DB	61	IleGluLeuHisGlyLysProIleGluValGluHisSerValProLysArgGlnArgIle	80
QY	313	AGGAAATTCAGATTCTGAACATCCCTCCTCACCTGCGAGGTGTTGGATGGACTT	372
DB	81	ArgLysLeuGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspSerLeu	100
QY	373	TGSGCTCAATATGGACAGTGGAGATGTGGAACAGTCAACACAGACACAGAACCGCC	432
DB	101	LeuValGlnTyrlleGlyValValGluSerCysGluGlnValAsnThrAspSerGluThrAla	120
QY	433	GTGTGCAACGTCACATATGCAACAGAGCAAGCAAAATAGCCATCGAGAGCTAAGC	492
DB	121	ValValAsnValThrTyrlleSerSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn	140
QY	493	GGGCATCAGTTTGAGAACTACTCCTTCAAGATTTCTACATCCCGGATGAAGAGGTGAGC	552
DB	141	GlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrlleProAspGluThrAla	160
QY	553	TCCCTTCGCCCTTCAG-----CGAGCCGAGCGT-----GGGACCATCTTCCCGG	600
DB	161	GlnGlnAsnProLeuGlnGlnProArgGlyArgGlyLeuGlyGlnArgGlySerSer	180

QY	601	GAGCAAGCCACGCCCCCTGGGGGCACTTCTCAGGCCACACAGATTGATTTCCCGTGGCG	660
DB	181	ArgGlnGly---SerProGlySerValSerLysGlnLysProCysAspLeuProLeuArg	199
QY	661	ATCTGTGTCCTCCACCCAGCTTTGTGTGCCATCATCGGAAGAGGGCTTGACCAATAAG	720
DB	200	LeuLeuValProThrGlnPheValGlyAlaIleIleGlyLysGluGlyValaThrIleArg	219
QY	721	AACATCACTAAGCAGACCCAGTCCCGGGTAGATATCCATAGAAAAGAACTCTCGAGCT	780
DB	220	AsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGlyAla	239
QY	781	CGAGAGAAGCTGTCAACCATCCACCCAGAGGGGACTTCTGAAGCATGCGCGATG	840
DB	240	AlaGluLysSerIleThrIleLeuSerThrProGluGlyThrSerAlaAlaCysLysSer	259
QY	841	ATTCTTGAATCATGTCAGAAAGAGCGCATGAGACCACCAACTAGCCGAAGAGATTCCTCTG	900
DB	260	IleLeuGluIleMetHisLysGluAlaGlnAspIleLysPheThrGluGluIleProLeu	279
QY	901	AAATCTTTGGCACACAATGGCTTGTGGAAAGACTGATTGGAAGAGAGGAGCAAAATTG	960
DB	280	LysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsnLeu	299
QY	961	AAGAAATTTGAACATGAAGAGGACCAAGATACAATCTCATCTTTTCAGAGATTGAGC	1020
DB	300	LysLysIleGluGlnAspThrAspThrLysIleThrIleSerProLeuGlnGluLeuThr	319
QY	1021	ATATACAAACCGGAAAGAACCATCATGTGAAGGCGCACAGTTGAGGCTGTGCCAGTCT	1080
DB	320	LeuTyraAsnProGluArgThrIleThrValLysGlyAsnValGluThrCysAlaLysAla	339
QY	1081	GAGATAGAGATTATGAAGAAGCTCGTGAGCCCTTTGAAAATGATATGCTGGTGTAAAC	1140
DB	340	GluGluGluIleMetLysLysIleArgGluSerTyrlleAsnAspIleAlaSerMetAsn	359
QY	1141	CAACAGCAATCTGATCCCGAGGTTGACCTCAGCGGCACTTGGCATCTTTTCAACAGGA	1200
DB	360	LeuGlnAlaHisLeuIleProGlyLeuAsnLeuAsnAlaLeuGlyLeuPhe-----	376
QY	1201	CTGTCCGTGCTATCTCCACACAGCGGCGCCCGCGAGCTCCCGCGCTGCCCTTACCAC	1260
DB	377	-----ProProThrSer-----GlyMetProProProThr-----	386
QY	1261	CCCTTCACTACCCACTCCGATACTTCTCCAGCCTGTACCCCATCACCAGTTTGGCCCG	1320
DB	387	-----SerGlyProProSerAlaMetThrPro-----Pro	396
QY	1321	TTCCCGCATCATCTCTTATCCAGAGCAGAGATTGTGAATCTTCTCATCCCAACCCAG	1380
DB	397	TyrProGlnPheGluGln---SerGluThrGluThrValHisLeuPheIleProAlaLeu	415
QY	1381	GCTGTGGCGGCATCATCGGAAGAGGCGCACACATCAACAGCTGGCGAGATTCCGCC	1440
DB	416	SerValGlyAlaIleIleGlyLysGlnGlyGlnHisIleLysGlnLeuSerArgPheAla	435
QY	1441	GGAGCCTCTATCAAGATTGCCCTTCGGAAGGCCAGACGTCAGCGAAAGATGCTCATC	1500
DB	436	GlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLysValArgMetValIle	455
QY	1501	ATCACCGGCGCCCGAGGCCAGTTCAGGCCCGGAGCGGATCTTTGGGAACACTGAAA	1560
DB	456	IleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleTyrlleGlyLysIleLys	475
QY	1561	GAGGAAACTCTTTTAAACCCCAAGAGAGTGAAGCTGGAAGCGCATATCAGAGTGGCC	1620
DB	476	GluGluAsnPheValSerProLysGluGluValLysLeuGluAlaHisIleArgValPro	495
QY	1621	TCTTCCACAGTGGCCGGGTGATTGGCAAGAGGTGGCAAGACCGTGAACGAACTCGAAGC	1680
DB	496	SerPheAlaAlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsn	515
QY	1681	TTAACAGTGCAGAGTCATCGTCCCTCGTGACCAAGCCGAGATGAAATGAGGAAGTG	1740

Db 416 SerValGlyAlaIleIleGlyLysGlnGlyGlnHisIleLysGlnLeuSerArgPheAla 435
QY 1441 GGAGCCTCTATCAAGTTTCCTCCCTGCGGAGGCCGACGCTCAGCGAAGAGTGTGTCATC 1500
Db GlyAlaSerIleLysIleAlaProAlaGluAlaProAspAlaLysValArgMetValIle 455
QY 1501 ATCACCAGGCGCCAGGAGCCAGATTCAAGGCCCGAGGACGGATCTTTGGGAACTGAAA 1560
Db 11eThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIleGlyLysIleLys 475
QY 1561 GAGGAAATCTTTAAACCCCAAGAAAGAGTGAAGCTGGAAGCGCATATCAGATGCC 1620
Db 476 GluGluAsnPheValSerProLysGluGluValLysLeuGluAlaHisIleArgValPro 495
QY 1621 TCTTCCACAGCTGGCGGTGATTGGCAAGGTGGCAGACCGTGAACGAATCGCAGAAC 1680
Db 496 SerPheAlaAlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsn 515
QY 1681 TPAACCACTGCAAGATCATCTGCTCGTGCACCAACGCCAGATCAAAATCAGGAAGTG 1740
Db 516 LeuSerSerAlaGluValValProArgAspGlnThrProAspGluAsnAspGlnVal 535
QY 1741 ATCTGCAAGATTATCGGCACCTCTTGTAGCCAGATGTCACAGCGCAAGATCAGGAAA 1800
Db 536 ValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGlnArgLysIleGlnGlu 555
QY 1801 ATTGTACACAGTGCAAGCAGCAGGAGCAGAAA 1833
Db 556 11eLeuThrGlnValLysGlnHisGlnGlnGln 566

RESULT 15

US-10-007-700-446
; Sequence 446, Application US/10007700
; Publication No. US20030064947A1

GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Li, Samuel X.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; APPLICANT: Retter, Marc W.
; APPLICANT: Durham, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Peckman, David W.
; APPLICANT: Cai, Feng
; APPLICANT: Fov, Teresa M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C17
; CURRENT APPLICATION NUMBER: US/10/007,700
; CURRENT FILING DATE: 2001-11-30
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 446
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-007-700-446

Alignment Scores:

Prod. No.: 2.67e-145 Length: 579
Score: 1957.50 Matches: 387
Percent Similarity: 79.19% Conservative: 81
Best Local Similarity: 65.48% Mismatches: 94
Query Match: 31.98% Indels: 29
Db: 12 Gaps: 8

US-09-270-437D-6 (1-3412) x US-10-007-700-446 (1-579)
QY 73 ATGAACAAGCTTTACATCGGAACTCTGAGCCCGCGCTCACCCGCCAGACCTCCGGCAG 132
Db 1 MetAsnLysLeuTyrIleGlyAsnLeuSerGluAsnAlaAlaProSerAspLeuGluSer 20
QY 133 CTCCTTTGGGACAGAAAGCTGCCCTGGCGGACAGGTCTCTGCTGAAGTCCGGCTACGCC 192
Db 21 11ePheLysAspAlaLysIleProValSerGlyProPheLeuValLysThrGlyTyrAla 40
QY 193 TTTCTGGGACTACCCGACGAGAACTCGGCATCCGCGCATCGAGACCTCTCGGTTAAA 252
Db 41 PheValAspCysProAspGluSerIlePheLeuLysAlaIleGluAlaLeuSerGlyLys 60
QY 253 GTGGAATTCATCGGAAATCATGGAAGTTGATTACTCAGTCTCTAAAAAGCTAAAGCAGC 312
Db 61 11eGluLeuHisGlyLysProIleGluValGluHisSerValProLysArgGlnArgIle 80
QY 313 AGGAAATTCAGATTTCGAAACATCCTCTCTACCTGCGAGTGGAGGTGTGATGAGCTT 372
Db 81 ArgLysLeuGlnIleArgAsnIleProHisLeuGlnTrpGluValLeuAspSerLeu 100
QY 373 TTGGCTCAATATGCGACAGTGGAGAAATGTGGAACAAGTCAACACAGACACAGAAACGCC 432
Db 101 LeuValGlnTyrGlyValValGluSerCysGluGlnValAsnThrAspSerGluThrAla 120
QY 433 GTTGTCAACGTCAATATGCAACAGAGAAGAAGCAAAATAGCCATCGAAGAAGCTAAAGC 492
Db 121 ValValAsnValThrTyrSerSerLysAspGlnAlaArgGlnAlaLeuAspLysLeuAsn 140
QY 493 GGGCATCAGTTTGAAGACTACTCTTCAAGATTCTCTACATCCCGGATGAGAGGTGAGC 552
Db 141 GlyPheGlnLeuGluAsnPheThrLeuLysValAlaTyrIleProAspGluThrAla 160
QY 553 TCCCTCTCGCCCTCAG-----CGAGCCAGAGCT-----GGGAGCACCCTCTCCCGG 600
Db 161 GlnGlnAsnProLeuGlnGlnProArgGlyArgArgGlyLeuGlyGlnArgGlySerSer 180
QY 601 GAGCAAGGCCACGCCCTTGGGGGCATTTCTCAGGCCACAGACAGATTGATTTCCCGTGGCG 660
Db 181 ArgGlnGly---SerProGlySerValSerLysGlnLysProCysAspLeuProLeuArg 199
QY 661 ATCTCTGCTCCCGCCAGCTTTGTTGGTCCCATCTCGGAAAGGAGGGCTTGACCATAAAG 720
Db 200 LeuLeuValProThrGlnPheValGlyAlaIleIleGlyLysGluGlyAlaThrIleArg 219
QY 721 AACATCACTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAAGAGAACTCTGGAGCT 780
Db 220 AsnIleThrLysGlnThrGlnSerLysIleAspValHisArgLysGluAsnAlaGlyAla 239
QY 781 GCAGAGAAGCTGTACCATCCATGCCACCCAGAGGGGACTTCTGAGCATGCGCATG 840
Db 240 AlaGluLysSerIleThrIleLeuSerThrProGluGlyThrSerAlaAlaCysLysSer 259
QY 841 ATCTCTGAATCATCGAAGAGCGCAGATGAGACCAACTAGCCGAGAGATTCCTCTG 900
Db 260 11eLeuGluIleMetHisLysGluAlaGlnAspIleLysPheThrGluGluIleProLeu 279
QY 901 AAAATCTTGGCACACAATGGCTTGTGGTGGAGACTGATTGGAAAAGAGGAGGAGAAATTG 960
Db 280 LysIleLeuAlaHisAsnAsnPheValGlyArgLeuIleGlyLysGluGlyArgAsnLeu 299
QY 961 AAGAAATTTGAACATGAACAGGACCAAGATCAACATCTCATCTTTGCGAGATTTCGAGC 1020
Db 300 LysLysIleGluGlnAspThrAspThrLysIleThrIleSerProLeuGlnLeuThr 319
QY 1021 ATATACACCGGNAAGAACCATCATCTGTGAAGGCGACAGTTGAGGCGCTGTGCCAGTCT 1080
Db 320 LeuTyrAsnProGluArgThrIleThrValLysGlyAsnValGluThrCysAlaLysAla 339
QY 1081 GAGATAGAGATTATGAAGAAGCTGCGTGGAGCCCTTTCAAAATGATATGCTGGCTGTTAAC 1140
Db 340 GluGluGluIleMetLysLysIleArgGluSerTyrGluAsnAspIleAlaSerMetAsn 359


```
QY 1141 CAACAGCCAACTCATCCAGGTTGAACCTCAGCGCATCTTGGCATCTTTCAACAGGA 1200
Db |||||:::|||||
360 LeuGlnAlaHisLeuIleProGlyLeuAsnLeuAlaLeuGlyLeuPhe----- 376
QY 1201 CTGTCCGTCTATCTCCACAGCAGGGGCCCGGAGCTCCCGCCCTGCCCTACCAAC 1260
Db |||||
377 -----ProThrSer-----GlyMetProProThr----- 386
QY 1261 CCTTCACTACCCGATCTCCGATCTCTCCAGCTGTACCCCATCACCAGTTTGGCCCG 1320
Db |||||
387 -----SerGlyProProSerAlaMetThrPro-----Pro 396
QY 1321 TTCCCGCATCATCTCTTATCCAGCAGAGGATTGTGAATCTCTTCATCCCAACCCAG 1380
Db ::::|
397 TyrProGlnPheGluGln---SerGluThrGluThrValHisLeuPheIleProAlaLeu 415
QY 1381 GCTGTGGGCGCCATCATCGGGAAGAGGGGCACACATCAAAACAGCTGGCGAGATTCGCC 1440
Db ::|||
416 SerValGlyAlaIleIleGlyGlnGlyGlnHisIleIleGlnLeuSerArgPheAla 435
QY 1441 GGAGCTCTATCAAGATTCCCTCGGGAAGGCCAGAGCTCAGCGAAAGGATGGTCATC 1500
Db |||||
436 GlyAlaSerIleIleIleAlaProAlaGluAlaProAspAlaLysValArgMetValIle 455
QY 1501 ATCAGCGGCGCCAGCGAGCCAGTTCAAGGCCCGAGCGGATCTTTGGAAACTGAAA 1560
Db |||||
456 IleThrGlyProProGlnAlaGlnPheLysAlaGlnGlyArgIleIleGlyIleLys 475
QY 1561 GAGGAAACTCTTTAAACCCCAAGAGAGTGAAGCTGGAAGCGCATATCAGAGTSCCC 1620
Db |||||
476 GluGluAsnPheValSerProLysGluGluValLysLeuGluAlaHisIleArgValPro 495
QY 1621 TCTTCCACAGCTGCGCGGTGATTGGCAAGGTGGCAAGACCGTGAAACGAACTGCAGAAC 1680
Db |||||
496 SerPheAlaAlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsn 515
QY 1681 TTAACCAAGTCAGAGTCTCGTCCCTCGTGACCAACGCCAGATGAAATAGGAAGTG 1740
Db |||||
516 LeuSerSerAlaGluValValProArgAspGlnThrProAspGluAsnAspGlnVal 535
QY 1741 ATCGTCAGAAATTATCGGGCACTTCTTTGCTAGCCAGACTGCACAGCGCAAGATCAGGGAA 1800
Db ::|||
536 ValValLysIleThrGlyHisPheTyrAlaCysGlnValAlaGlnArgLysIleGlnGlu 555
QY 1801 ATTGTACAACAGTGAAGCAGCAGGACGAGAAA 1833
Db |||||
556 IleLeuThrGlnValLysGlnHisGlnGlnGln 566
```

Search completed: July 16, 2004, 10:57:42

Job time : 357 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: July 16, 2004, 10:47:12 ; Search time 149 Seconds
(without alignments)
12940.296 Million cell updates/sec

Title: US-09-270-437D-6

Perfect score: 1120

Sequence: 1 ggcagcgaggagcgagga.....aaccttgaaaaatgtttattt 3412

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1586107 seqs, 282547505 residues

Word size: 1

Total number of hits satisfying chosen parameters: 2987630

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

--MODEL=frame_n2p.model -DEV=xlp
-Q=/cgn2/1/USPRO.spool.p/US09270437/runat 16072004.113126.13935/app query.fasta.1.3591
-DB=A Geneseq 29Jan04 -QFMT=fastan -SUFFIX=Oligo.rag -MINMATCH=0.1 -LoopCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS-human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEA8IZE=500 -MINLEN=0 -MAXLEN=2000000000
-USFR=US09270437.cgn_1_1_253_orunat 16072004.113126.13935 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : A Geneseq 29Jan04.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	603	53.8	614	4	ABG06794 Novel hum
2	511	45.6	620	4	AAU16163 Human nov
3	511	45.6	620	6	ABU5232 Human nov
4	378	33.8	594	4	ABG06795 Novel hum
5	357	31.9	556	5	ABG96346 Human ova
6	160	14.3	171	4	AAU16166 Human nov
7	160	14.3	171	6	ABU5235 Human nov
8	97	8.7	97	3	AAG03261 Human sec
9	75	6.7	192	4	AAU16164 Human nov
10	75	6.7	192	6	ABU5233 Human nov

11	66	5.9	171	4	AAU16583
12	66	5.9	171	6	ABU55652
13	44	3.9	555	6	ABU89799
14	32	2.9	187	6	AAO23971
15	20	1.8	47	4	AAU16581
16	20	1.8	47	6	ABU55650
17	19	1.7	148	4	ABG21962
18	19	1.7	250	4	AAU16579
19	19	1.7	250	6	ABU55648
20	19	1.7	261	4	AAU16161
21	19	1.7	261	6	ABU55230
22	19	1.7	319	4	AAU53826
23	19	1.7	577	2	AAU30649
24	19	1.7	579	3	AAU11328
25	19	1.7	579	3	AAU11365
26	19	1.7	579	5	ABU74960
27	19	1.7	579	5	ABU75053
28	19	1.7	579	5	ABU74997
29	19	1.7	579	5	ABU75054
30	19	1.7	579	5	ABU61917
31	19	1.7	579	5	ABU61974
32	19	1.7	579	5	ABU61880
33	19	1.7	579	5	ABU61973
34	19	1.7	579	7	ADA28536
35	19	1.7	579	7	ADA28539
36	19	1.7	579	7	ADA28438
37	19	1.7	579	7	ADA28266
38	19	1.7	579	7	ADD14066
39	19	1.7	579	7	ADU53471
40	19	1.7	586	5	ABU75048
41	19	1.7	586	5	ABU61968
42	19	1.7	586	7	ADA28517
43	18	1.6	619	4	ABG21963
44	17	1.5	583	4	ABG12592
45	15	1.3	20	5	ABU75062

ALIGNMENTS

RESULT 1
ABG06794

ID ABG06794 standard; protein; 614 AA.

AC ABG06794;

XX

DT 13-FEB-2002 (first entry)

XX

DE Novel human diagnostic protein #6785.

XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder.

XX

OS Homo sapiens.

XX

PN WO200175067-A2.

XX

PD 11-OCT-2001.

XX

PF 30-MAR-2001; 2001WO-US008631.

XX

PR 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Drmanac RT, Liu C, Tang YT;

XX

DR WPI; 2001-639362/73.

DR N-PSDB; AAS70981.

XX

PT New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity.

XX Claim 20; SEQ ID NO 37153; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG0010-ABG0377 represent novel human diagnostic amino acid sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 614 AA;

Alignment Scores:

Pred. No.: 0 Length: 614
Score: 603.00 Matches: 603
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 53.84% Indels: 0
DB: 4 Gaps: 0

US-09-270-437D-6 (1-3412) x ABG06794 (1-614)

QY 58 GGGAGAGACGGATGATGAACAGCTTTACATCGGGAACCTGAGCCCGCGCTCACGGC 117
DB 12 GlyLysArgArgMetMetAsnLysLeuTyrIleGlyAsnLysSerProAlaValThrAla 31
QY 118 GACGACCTCCGACGCTCTTGGGACAGGAGCTGCCCTCGCGGACAGCTCTCTGCTG 177
DB 32 AspAspLeuArgGlnLeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLeu 51
QY 178 AAGTCGGCTACGCTCTGCTGACTACCCGACACAGAACTGGGCGCATCCGCGCATCGAG 237
DB 52 LysSerGlyTyrAlaPheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGlu 71
QY 238 ACCCTCTCGGTAAAGTGAATTGCATGGGAAATCATGGAAGTTGATTACTCAGTCTCT 297
DB 72 ThrLeuSerGlyLysValGluLeuHisGlyLysIleMetGluValAspTyrSerValSer 91
QY 298 AAAAGACTAAGACGAGGAGAAATTCAGATTCCAAACATCCCTCCTCACCTCGAGTGGAG 357
DB 92 LysLysLeuArgSerArgLysIleGlnIleArgAsnIleProProHisLeuGlnTrpGlu 111
QY 358 GTGTTGGATGGACTTTTGGCTCAATATGGGACAGTGGAGAAATGTGAAACAAGTCACACA 417
DB 112 ValLeuAspGlyLeuLeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThr 131
QY 418 GACAGAAACCGCGTCTGACAGTCACATATGCACAGACAGGAGGAGGAGGAGGAGGAGG 477
DB 132 AspThrGluThrAlaValValAsnValThrTyrAlaThrArgGlnGluAlaLysIleAla 151
QY 478 ATGGAGAAGCTAAGCGGGCATCAGTTTGAGAACTACTCTCTTCAAGATTTCTACATCCCG 537
DB 152 MetGluLysLeuSerGlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIlePro 171
QY 538 GATGAAGAGGTGAGTCCCTTCGCGCCCTCAGCAGGCCAGCGTGGGACCACTTCTCC 597
DB 172 AspGluGluValSerSerProSerProGlnArgAlaGlnArgGlyAspHisSerSer 191

QY 598 CGGAGACAGGACCAACCGCCCTGGGGGCACTTCTCAGGCCACAGATTGATTTCCGCTG 657
DB 192 ArgGluGlnGlyHisAlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeu 211
QY 658 CGGATCTGGTCCCGCCACCCAGTGTGGTCCCATCATCGAAAGAGGGCTTGACCATATA 717
DB 212 ArgIleLeuValProThrGlnPheValGlyAlaIleIleGlyLysGluGlyLeuThrIle 231
QY 718 AAGAACATCATTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAAAGAACTCTCGA 777
DB 232 LysAsnIleThrLysGlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGly 251
QY 778 GCTGAGAGAACCTGTCTACCATTCATCCACCCAGAGGGAGCTTCTGAGACATCCGCG 837
DB 252 AlaAlaGluLysProValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArg 271
QY 838 ATGATTTCTTGAATCATGCAGAAAGAGGACAGATGAGCAAACTAGCCGAGAGATTCT 897
DB 272 MetIleLeuGluIleMetGlnLysGluAlaAspGluThrLysLeuAlaGluGluLeuPro 291
QY 898 CTGAAAATCTTGGCACACAATGGCTGGTGGAGACCTGATTGGAAGAGAGGAGAAAT 957
DB 292 LeuLysIleLeuAlaHisAsnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsn 311
QY 958 TTGAAGAAATTTGAACATGAAACAGGACAGGACCAAGATAACAATCTCATCTTTCAGGATT 1017
DB 312 LeuLysIleGluIleGluIleThrGlyThrLysIleThrIleSerSerLeuGlnAspLeu 331
QY 1018 AGCATATACAAACCCGGAAGAACCATCATCTGTGAAGGACAGATTGAGCCCTGTCCAGT 1077
DB 332 SerIleTyrAsnProGluArgThrIleThrValLysGlyThrValGluAlaCysAlaSer 351
QY 1078 GCTGAGATAGATATTAAGAGAGCTGCTGAGGCTTTGAAATGATATGCTGCTGTT 1137
DB 352 AlaGluIleGluIleMetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaVal 371
QY 1138 AACCAACAGCAATCTGATCCCGGTTGAACCTCAGCGACCTGAGGACCTTTTCAACA 1197
DB 372 AsnGlnGlnAlaAsnLeuIleProGlyLeuAsnLeuSerAlaLeuGlyIlePheSerThr 391
QY 1198 GGACTGTGCTGCTATCTCCACAGAGGCGCCCGGAGCTCCCGCTCCCGCTCCCGCTAC 1257
DB 392 GlyLeuSerValLeuSerProProAlaGlyProArgGlyAlaProProAlaAlaProTyr 411
QY 1258 CACCCCTTACTACCCACTCCGATCTTCTCCAGCTGTACCCCTACCCCTACCCAGTTGCG 1317
DB 412 HisProPheThrThrHisSerGlyTyrPheSerSerLeuTyrProHisHisGlnPheGly 431
QY 1318 CGTTTCCCGCATCATCTCTTATCCAGACGAGAGATTGTAATCTCTTCATCCCAAC 1377
DB 432 ProPheProHisHisSerTyrProGluGlnGluIleValAsnLeuPheIleProThr 451
QY 1378 CAGGCTGTGGGCGCATCATCGGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1437
DB 452 GlnAlaValGlyAlaIleIleGlyLysLysGlyAlaHisIleLysGlnLeuAlaArgPhe 471
QY 1438 GCGGAGCTCTTCAAGATTCGCCCTCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1497
DB 472 AlaGlyAlaSerIleLysIleAlaProAlaGluGlyProAspValSerGluArgMetVal 491
QY 1498 ATCATCACCGGCGCCAGGAGCCAGTTCAAGGCGGAGGAGGAGGAGGAGGAGGAGGAGG 1557
DB 492 IleIleThrGlyProProGluAlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeu 511
QY 1558 AAAAGAGAAATCTTCTTAAACCCCAAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1617
DB 512 LysGluGluAsnPheAsnProLysGluGluValLysLeuGluAlaHisIleArgVal 531
QY 1618 CCTCTTCCACAGCTGGCGGCTGATTGGCAAGGTGGCAAGACCGTGAACCACTTCAG 1677
DB 532 ProSerSerThrAlaGlyArgValIleGlyLysGlyLysGlyLysValAsnGluLeuGln 551

PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX

PI Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-488783/53.

DR N-PSDB; AAS26150.

XX New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives.

XX Claim 11; SEQ ID NO 1116; 980pp; English.

XX The invention relates to isolated nucleic acid molecules and their
CC encoded secreted proteins. The nucleic acids and proteins are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. Antibodies to the proteins can also be used in
CC alleviating symptoms associated with the disorders and in diagnostic
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays
CC (ELISA). Disorders which are diagnosed or treated include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
CC and many other disorders listed in the specification. The polypeptides
CC can also be used to aid wound healing and epithelial cell proliferation,
CC to prevent skin aging due to sunburn, to maintain organs tissues,
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence represents a novel secreted protein of the invention. Note: The
CC sequence data for this patent did not form part of the printed

Alignment Scores:

Pred. No.: 0 Length: 620
Score: 511.00 Matches: 511
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 45.62% Indels: 0
DB: 4 Gaps: 0

US-09-270-437D-6 (1-3412) x AAUI6163 (1-620)

QY 70 ATGATGAACAAAGCTTTACATCGGAAACCTGAGCCGCCGCTCACCGCGAGCACTCCGG 129
|||
Db 22 MetMetAsnLysLeuTyrIleGlyAsnLeuSerProAlaValThrAlaAspLeuArg 41
|||
QY 130 CAGCTCTTTGGGACAGAGAGCTGCCCTGGCGGACAGGTCTCTGCTCAAGTCGGCTAC 189
|||
Db 42 GlnLeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLeuLysSerGlyTyr 61
|||
QY 190 GCCTTCGTGGACTACCCCGACAGAACTGGGCGCATCCGCGCATCGAGACCTCTCGGT 249
|||
Db 62 AlaPheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGluThrLeuSerGly 81
|||
QY 250 AAAGTGAATTCCATGCGAAATCATGGAAGTTGTAATCTCAGTCTCTAAAAAGTAAAG 309
|||
Db 82 LysValGluLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLysLeuArg 101
|||
QY 310 AGCAGGAAATTCAGATTCGAAACATCCCTCCTCACCTGCGAGTGGAGGTGTGGATGGA 369
|||
Db 102 SerArgLysIleGlnIleArgAsnIleProHisLeuGlnTrpGluValLeuAspGly 121
|||
QY 370 CTTTGGCTCAATATGGGACAGTGGGAATGTGGAACAAGTCAACACACACAGAAACC 429
|||
Db 122 LeuLeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThrAspThrGluThr 141
|||
QY 430 GCCGTTGTCAAGTCAATATGCAATATGCAACAGAGAGAGCAAAATAGCCATGAGAGCTA 489
|||
Db 142 AlaValValAsnValThrTyrAlaThrArgGluGluAlaLysIleAlaMetGluLysLeu 161
|||
QY 490 AGCGGSCATCAGTTTGAGAACTACTCTTCAAGATTTCTACATCCCGATGAAGAGGTG 549
|||
Db 162 SerGlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIleProAspGluGluVal 181
|||
QY 550 AGCTCCCTTCGCCCTCAGCAGCCCGCAGCGTGGGACCACTCTTCCCGGAGCAGCAGC 609
|||
Db 182 SerSerProSerProProGlnArgAlaGlnArgGlyAspHisSerSerArgGluGlnGly 201
|||
QY 610 CAGGCCCTTGGGGGACCTTCTCAGGCCACAGATTTGTTCCCGTCGGGATCCTGTGTC 669
|||
Db 202 HisAlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeuVal 221
|||
QY 670 CCCACCCAGTTTGTGGTCCCATCATCGGAAAGGAGGGCTTGACCATAAAGACATCACT 729
|||
Db 222 ProThrGlnPheValGlyAlaIleIleGlyLysGluGlyLeuThrIleLysAsnIleThr 241
|||
QY 730 AAGCAGACCCAGTCCCGGTAGATATCCATAGAAAGAGAACTCTGGAGCTGCAGACAG 789
|||
Db 242 LysGlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGlyAlaAlaGluLys 261
|||
QY 790 CCTGTCCACATCCATCCGCCCGGAGCTTCTGAAGCATGCCGCGATGATTTCTTGA 849
|||
Db 262 ProValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeuGlu 281
|||
QY 850 ATCATCAGAAAGAGGACAGATGAGACCAACTAGCCGAAAGAGATTCCTCTGAAATCTTG 909
|||
Db 282 IleMetGlnLysGluAlaAspGluThrLysLeuAlaGluGluIleProLeuLysIleLeu 301
|||
QY 910 GCACACAATGGCTTGGTGGAGACTGATTTGAAAGAGGCGCAAAATTTGAAGAAATTT 969
|||
Db 302 AlaHisAsnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLysLysIle 321
|||
QY 970 GAACATGAACACGGGACCAAGATAACAATCTCATCTTTGAGGATTTGACATATACAAC 1029
|||
Db 322 GluHisGluThrGlyThrLysIleThrIleSerSerSerLeuGlnAspLeuSerIleTyrAsn 341
|||

QY 1030 CCGGAAAGAACCATCACTGCTGAAGGCGACAGCTTGAGGCGCTGTGCAGTGTGAGATAGAG 1089
Db 342 ProGluArgThrIleThrValLysGlyThrValGluAlaCysAlaSerAlaGluIleGlu 361
QY 1090 ATTATGAAGAAGCTCGTGAGGCGCTTTGAAATGATATGCTGGCTGTAAACCAACAGCC 1149
Db 362 IleMetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaValAsnGlnGlnAla 381
QY 1150 AATCTGATCCAGGTTGAACCTCAGCGCACTTGGCATCTTTCAACAGGACTGTCGCTG 1209
Db 382 AsnLeuIleProGlyLeuAsnLeuSerAlaLeuGlyIlePheSerThrGlyLeuSerVal 401
QY 1210 CTATCTCCACAGAGGCGCGGAGCTCCCGCGCTCCCGCATCACCATCACCATCTTGGCCGCTTCCCGCAT 1329
Db 422 ThrHisSerGlyTyPheSerSerLeuTyProHisGlnPheGlyProPheProHis 441
QY 1330 CATCACTTTATCCAGACGAGAGATTGTGAATCTTTCATCCCAACCCAGGCTGTGGGC 1389
Db 442 HisHisSerTyProGluGlnGluIleValAsnLeuPheIleProThrGlnAlaValGly 461
QY 1390 GCCATCATCGGGAAGAAGGGGCGCACATCAACAGCTGGCGAGATTCCGCGGAGCTCT 1449
Db 462 AlaIleIleGlyLysLysGlyAlaHisIleLysGlnLeuAlaArgPheAlaGlyAlaSer 481
QY 1450 ATCAAGATTGCCCTCGGGAAGGCCCGACGTCAGCGAAGGATGCTCATCATCACCGGG 1509
Db 482 IleLysIleAlaProAlaGluGlyProAspValSerGluArgMetValIleIleThrGly 501
QY 1510 CCACCGAAGCCAGTTCAAGGCCCGAGCGGATCTTTGGGAAACTGAAAGAGGAAAC 1569
Db 502 ProProGluAlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluGluAsn 521
QY 1570 TTCTTTAACCCCAAGAGAGTCAAGCTGGA 1602
Db 522 PhePheAsnProLysGluGluValLysLeuGlu 532
RESULT 3
ID ABU55232 standard; protein; 620 AA.
XX AC ABU55232;
XX AC
XX AC
DT 18-MAR-2003 (first entry)
XX DE
XX DE
XX Human novel polypeptide #319.
KW Human; neural disorder; immune system disorder; renal disorder;
KW muscular disorder; respiratory disease; reproductive disorder;
KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
KW hyperproliferative disorder; inflammatory disease; allergic reaction;
KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
KW haemostatic; antiarteriosclerotic.
XX OS
XX Homo sapiens.
XX PN US2002132753-A1.
XX PD
XX PD
XX 19-SEP-2002.
XX PF 17-JAN-2001; 2001US-00764864.
XX PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214886P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX PI Rosen CA, Ruben SM, Barash SC;
XX WPI; 2003-147444/14.
XX N-PSDB; ABX73491.
XX New polypeptides and nucleic acids, useful in gene therapy for treating,
PT inhibiting or preventing e.g. neural, immune system, muscular,
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
XX renal disorders.
XX Claim 11; SEQ ID NO 1116; 402pp; English.
XX The invention relates to human novel polypeptides and their associated
CC polynucleotides. The polypeptides and polynucleotides are useful in gene
CC therapy for treating, inhibiting or preventing neural disorders, immune
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
CC appendicitis), allergic reactions and conditions (e.g. asthma), blood

CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
 CC infarction) and cancerous diseases. Sequences ABU54914-ABU55699 and
 CC ABU55748 represent human novel polypeptides of the invention
 XX
 SQ Sequence 620 AA;

Alignment Scores:

Pred. No.: 0 Length: 620
 Score: 511.00 Matches: 511
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 45.62% Indels: 0
 DB: 6 Gaps: 0

US-09-270-437D-6 (1-3412) x ABU55232 (1-620)

QY	70	ATGATGAACAAGCTTTACATCGGAACTGAGCCCGCGCTCACCGCGACGACCTCCGG	129
Db	22	MetMetAsnLysLeuTyriLeGlyAsnLeuSerProAlaValThrAlaAspAspLeuArg	41
QY	130	CAGCTCTTTGGGACAGGAAGCTGCCCTGGCGGACAGGTCTCTGAGTCCGGCTAC	189
Db	42	GlnLeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLysSerGlyTyr	61
QY	190	GCCTTCGTGGACTACCCCGACAGAACTGGGCGCATCCGCGCCATCGAGACCTCTCGGT	249
Db	62	AlaPheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGlnThrLeuSerGly	81
QY	250	AAAGTGGAAATTCATGGGAAATCATGGAAGTTGATTCTCAGTCTCTAAAAGCTAAG	309
Db	82	LysValGluLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLysLeuArg	101
QY	310	AGCAGAAATTCAGATTCGAAATCCCTCTCTACCTCAGTGGAGGTGTGATGGA	369
Db	102	SerArgLysIleGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspGly	121
QY	370	CTTTGGCTCAATATGGACAGTGGAGATGTGGAACTGAACTCAACACACACAGAAAC	429
Db	122	LeuLeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThrAspThrGluThr	141
QY	430	GCCGTGTCAACGTCACATATGACACAGAGAGAGAGCAAAATATGCCATGGAGAGCTA	489
Db	142	AlaValValAsnValThrTyrAlaThrArgGluGluAlaLysIleAlaMetGluLysLeu	161
QY	490	ACGGGCACTAGTTTGGACAACTACTCTTCAAGATTCTCTACATCCCGATGAGAGGTG	549
Db	162	SerGlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIleProAspGluGluVal	181
QY	550	AGCTCCCTTCGCCCTCAGCGAGCCCGAGCTGGGACCACTCTTCCCGGAGCAAGGC	609
Db	182	SerSerProSerProGlnArgAlaGlnArgGlyAspHisSerSerArgGluGlnGly	201
QY	610	CACGCCCTTGGGGCACTTCTCAGGCCAGACAGATTGATTTCCGCTGGGATCTCTGTC	669
Db	202	HisAlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeuVal	221
QY	670	CCACCCAGTTCTTGGTCCATATCGGAAAGAGGGCTTGACCAATAAGACATCACT	729
Db	222	ProThrGlnPheValGlyAlaIleIleGlyLysGluGlyLeuThrIleLysAsnIleThr	241
QY	730	AAGCAGACCCAGTCCCGGGTAGATATCCATAGAAAAGAGAACTCTGGAGCTGCAGAGAAG	789
Db	242	LysGlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGlyAlaAlaGluLys	261
QY	790	CTGTGACCATTCATGCCACCCAGAGGGGACTTCTGAGCATGCCGATGTTCTTGAA	849
Db	262	ProValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeuGlu	281
QY	850	ATCATCAGAAAGAGGACAGATGAGCCAACTAGCCGAAAGAGATTCTCTGAAAATCTTG	909
Db	282	IleMetGlnLysGluAlaAspGluThrLysLeuAlaGluGluIleProLeuLysIleLeu	301
QY	910	GCACACAATGGCTTGGTGGAGACTGATTGGAAAAGAGGACAGAAATTTGAAGAAAT	969

Db	302	AlaHisandGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLysLysIle	321
QY	970	GAAATGAAACAGGACCAAGATAAATCTCTATCTTTGAGGATTTGAGATATACAC	1029
Db	322	GluHisGluThrGlyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsn	341
QY	1030	CCGAAAGAACCATCACTGTGAAGGACACAGCTTCCAGGCTGTCCAGTGTGAGATAG	1089
Db	342	ProGluArgThrIleThrValLysGlyThrValGluAlaCysAlaSerAlaGluIleGlu	361
QY	1090	ATTATGAAGAGCTCGCTGAGGCCCTTTGAAAATGATATGCTGGCTGTGTTAAACAAAC	1149
Db	362	IleMetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaValAsnGlnAla	381
QY	1150	AATCTGATCCAGGGTTGAACCTCAGCGCATCTTGGCATCTTTCAACAGGACTGCGGTG	1209
Db	382	AsnLeuIleProGlyLeuAsnLeuSerAlaLeuGlyIlePheSerThrGlyLeuSerVal	401
QY	1210	CTATCTCCACAGAGGCCCGCGAGCTCCCGCGCTGCCCTACACCCCTTCACT	1269
Db	402	LeuSerProProAlaGlyProArgGlyAlaProProAlaAlaProTyrHisProPheThr	421
QY	1270	ACCACTCCGATACTCTCCAGCCTGTACCCCATCACAGTTTGGCCCGTTCGCGCAT	1329
Db	422	ThrHisSerGlyTyrPheSerSerLeuTyrProHisGlnPheGlyProPheProHis	441
QY	1330	CATCACTTTATCCAGAGCAGAGATTTGTAATCTCTTCATCCCAACCCAGGCTGTGGC	1389
Db	442	HisHisSerTyrProGluGlnGluIleValAsnLeuPheIleProThrGlnAlaValGly	461
QY	1390	GCATCATCGGGAAGAGGGGACACATCAACAGCTGGCGAGATTCCCGGAGCCTCT	1449
Db	462	AlaIleIleGlyLysLysGlyAlaHisIleLysGlnLeuAlaArgPheAlaGlyAlaSer	481
QY	1450	ATCAAGATTGCCCTCGCGAAGCCCGACAGTCAAGGAAAGGATGTTATCATCACC	1509
Db	482	IleLysIleAlaProAlaGluGlyProAspValSerGluArgMetValIleIleThrGly	501
QY	1510	CCACCGGAGCCAGTTCAGGCCCCAGGACCGATCTTGGCAACTGAAAGAGGAAAC	1569
Db	502	ProProGluAlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluGluAsn	521
QY	1570	TTCTTTAAACCCCAAGAGAGTGAAGCTGGAA	1602
Db	522	PhePheAsnProLysGluGluValLysLeuGlu	532
RESULT 4			
ID	ABG06795		
XX	ABG06795	standard; protein; 594 AA.	
AC	ABG06795;		
XX	DT	(first entry)	
XX	DE	Novel human diagnostic protein #6786.	
XX	OS	Homo sapiens.	
XX	PN	WO200175067-A2.	
XX	PD	11-OCT-2001.	
XX	PF	30-MAR-2001; 2001WO-US008631.	
XX	PR	31-MAR-2000; 2000US-00540217.	
XX	PR	23-AUG-2000; 2000US-00649167.	
XX	PA	(HYSE-) HYSEQ INC.	
XX	XX	Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.	

PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS70982.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PT Claim 20; SEQ ID NO 37154; 103pp; English.
 PS
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
 CC patent did not appear in the invention. Note: The sequence data for this
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 594 AA;
 Alignment Scores:
 Pred. No.: 0 Length: 594
 Score: 378.00 Matches: 378
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 33.75% Indels: 0
 DB: 4 Gaps: 0
 US-09-270-437D-6 (1-3412) x ABG06795 (1-594)
 QY 7 GAGAGGCGGAGGAGCGCGGTACCGGCGGGGGAGCGCGGCTCTCGGGAGAGA 66
 Db 18 GlyGlyGlyGluGluArgArgValProGlyArgGlySerArgGlyLeuSerGlyLysArg 37
 QY 67 CGGATGATGAACAGCTTTACATCGGGAACTGAGCCCGCCCGCTACCGCGGACGACCTC 126
 Db 38 ArgMetMetAsnLysLeuTyrlleGlyAsnLeuSerProAlaValThrAlaAspLeu 57
 QY 127 CGCAGCTCTTTGGGACAGGAAGCTGCCCTCGGGGACAGCTCTCGTGAAGTCCGGC 186
 Db 58 ArgGlnLeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLysSerGly 77
 QY 187 TAGCCCTCTGTGACTACCCGACCGAAGCTGGGCCATCGCGCCATCGAGACCTCTCG 246
 Db 78 TyrAlaPheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGluThrLeuSer 97
 QY 247 GGTAAAGTGAATTGCATGGGAAATCAGGAAGTTGATTCTCAGTCTCTAAAGCTA 306
 Db 98 GlyLysValGluLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLysLeu 117
 QY 307 AGGACGAGGAATAATTCAGATTCCGAAACATCCCTCTCAGCTGAGGGAGGTGTCGAT 366
 Db 118 ArgSerArgLysIleGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAsp 137
 QY 367 GGACTTTGGCTCAATATATGGACAGTGGGAATGTGGAACAGTCAACACAGACAGAA 426
 Db 138 GlyLeuLeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThrAspThrGlu 157

QY 427 ACCGCGGTTGTCAAGTCACATATGCAACAAGAGAGCAAAATAATAGCCATCGAGAAG 486
 Db 158 ThrAlaValValAsnValThrTyraIaThrArgGluGluAlaLysIleAlaMetGluLys 177
 QY 487 CTAAAGCGGCGCATCAGTTTGTAGAACTACTCTTCAAGATTTCCTACATCCCGATCAAGAG 546
 Db 178 LeuSerGlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIleProAspGluGlu 197
 QY 547 GTGAGCTCCCTTCGCGCCCTCAGCGAGCCCGAGCGTGGGACCACTCTCTCCCGGAGCAA 606
 Db 198 ValSerSerProSerProGlnArgAlaGlnArgGlyAspHisSerSerArgGluGln 217
 QY 607 GGCCACGCGCCCTGGGGGCACTTCTCAGCCAGACAGATTGATTTCCGCTCGGATCTCTG 666
 Db 218 GlyHisAlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeu 237
 QY 667 GTCCCCACCCAGTTTGTGTGTCATCATCGAAAGGAGGCGCTTGACCATAAAGAAACATC 726
 Db 238 ValProThrGlnPheValGlyAlaIleIleGlyLysGluGlyLeuThrIleLysAsnIle 257
 QY 727 ACTAAGCAGACCCAGTCCCGGTAGATATCATAGAAAAGAGAACTCTGGAGCTGCGAG 786
 Db 258 ThrLysGlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGlyAlaAlaGlu 277
 QY 787 AAGCCTCTCACCATCCATGCCACCCAGAGGGGACTTCTGAAGCATGCCCATGATCTT 846
 Db 278 LysProValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeu 297
 QY 847 GAATCATGCAAGAAAGGAGGATGAGACCAAACTAGCCGAAAGAGATTCTCTGAAAATC 906
 Db 298 GluIleMetGlnLysGluAlaAspGluThrLysLeuAlaGluGluIleProLeuLysIle 317
 QY 907 TTGSCACACAAATGCTGTGGTGGNAGACTATTGGAAAAGAGAGCAAAATTTGAAGAA 966
 Db 318 LeuAlaHisAsnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLysLys 337
 QY 967 ATTGAACATGAAACAGGAGCAAGATAACAATCTCATCTTTGAGGAGTTTGACATATAC 1026
 Db 338 IleGluHisGluThrGlyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyr 357
 QY 1027 AACCCGGAAGAACCATCTACTGTGAAGGGCACAGTGTGAGCCCTGTGCGAGTGTGAGATA 1086
 Db 358 AsnProGluArgThrIleThrValLysGlyThrValGluAlaCysAlaSerAlaGluIle 377
 QY 1087 GACATTATGAAGAGCTGCTGAGGCGCTTTGAAAATGATATGCTGCTGTTAAC 1140
 Db 378 GluIleMetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaValAsn 395
 RESULT 5
 ABG96346
 ID ABG96346 standard; protein; 556 AA.
 XX
 AC ABG96346;
 XX
 DT 11-DEC-2002 (first entry)
 XX
 DE Human ovarian cancer marker M452.
 XX
 KW Human; ovarian cancer; marker; cancer; familial history; brain disorder;
 KW central nervous system disorder; bacterial meningitis; viral meningitis;
 KW Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;
 KW brain herniation; inflammation; encephalitis; testicular disorder;
 KW non-tuberculous granulomatous orchitis; connective tissue disorder;
 KW heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;
 KW histological type; carcinogenic; ovarian cancer marker.
 XX
 OS Homo sapiens.
 XX
 PN WO200271928-A2.
 XX
 PD 19-SEP-2002.
 XX

PF 14-MAR-2002; 2002WO-US007826.
XX
PR 14-MAR-2001; 2001US-0276025P.
PR 14-MAR-2001; 2001US-0276026P.
PR 10-AUG-2001; 2001US-0311732P.
PR 19-SEP-2001; 2001US-0323580P.
PR 26-SEP-2001; 2001US-0324967P.
PR 26-SEP-2001; 2001US-0325102P.
PR 26-SEP-2001; 2001US-0325149P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
XX Monahan JE, Gannavarapu M, Hoersch S, Kamatkar S, Kovatis SG;
PI Meyers RE, Morrisey MP, Olandt PU, Sen A, Vieby PO, Mills GB;
PI Bast RC, Lu K, Schmandt RE, Zhao X, Glatt K;
XX
DR WPI; 2002-723277/78.
DR N-PSDB; ABS76442.
XX
PT Assessing whether a patient is afflicted with ovarian cancer, useful in
PT assessing the stage or progression of the disease, comprises comparing
PT the expression level of a cancer marker in a sample from a patient and
PT from a non cancer patient.
XX
XX Disclosure; Page 263-264; 481pp; English.
XX
CC The present invention relates to a new method for assessing whether a
CC patient is afflicted with ovarian cancer. The method involves comparing
CC the expression level of a marker in a patient sample and the normal level
CC of expression of the marker in a control non-ovarian cancer sample, where
CC the marker is selected from 363 cancer markers described in the
CC specification. The method of the invention is useful in diagnosing or
CC characterising cancer, in detecting the presence of cancer as early as
CC possible, and the recurrence of ovarian cancer. The method may also be of
CC particular use with patients having an enhanced risk of developing
CC ovarian cancer (e.g. patients having a familial history of ovarian
CC cancer). The cancer markers may be used in the management and treatment
CC of e.g. brain and central nervous system disorders (e.g. bacterial and
CC viral meningitis, Alzheimer's disease or Parkinson's disease), brain
CC disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),
CC inflammations (e.g. bacterial or viral meningitis or encephalitis),
CC testicular disorders (e.g. nonbacterial granulomatous orchitis),
CC connective tissue disorders, or heart disorders (e.g. ischaemic heart
CC disease or atherosclerosis). The compositions and methods may also be
CC used in assessing the histological type of neoplasm associated with
CC ovarian cancer, monitoring the progression of ovarian cancer, determining
CC whether ovarian cancer has metastasized or is likely to metastasize,
CC selecting a composition for inhibiting ovarian cancer, assessing the
CC ovarian carcinogenic potential of a compound, or inhibiting ovarian
CC cancer or at risk of developing ovarian cancer. The present amino acid
CC sequence represents one of the ovarian cancer markers described in the
CC invention
XX
SQ Sequence 556 AA;

Alignment Scores:
Pred. No.: 0 Length: 556
Score: 357.00 Matches: 357
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 31.87% Indels: 0
DB: 5 Gaps: 0

US-09-270-437D-6 (1-3412) x ABS96346 (1-556)

QY 70 ATGATGACAAAGCTTTTACATCGGAACTGAGCCCGCCGCTCCCGCGACGACTCCGG 129
Db 1 MetMetAsnLysLeuTyrIleGlyAsnLeuSerProAlaValThrAlaAspLeuArg 20
QY 130 CAGCTCTTTGGGACAGAGCTGCCCTGGCGGACAGGTCTCTGCTGAAGTCCGGGTAC 189
Db 21 GlnLeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLeuLysSerGlyTyr 40

QY 190 GCCTTCGTGGACTACCCGACCCAGAACTGGGCCCATCGCGCCATCGAGACCTCTCCGGT 249
Db |||||
41 AlaPheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGluThrLeuSerGly 60
QY 250 AAAGTGAATTCATGCGGAAATCATGGAAGTTGATTCTCAGTCTCTTAAAGAACTAAGG 309
Db |||||
61 LysValGluLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLysLeuArg 80
QY 310 AGCAGGAAATTCAGATTCGAAACATCCCTCCTCACCTGCAGTGGAGGTGTGATGATGGA 369
Db |||||
81 SerArgLysIleGlnIleArgAsnIleProProHisLueGlnTrpGluValLeuAspGly 100
QY 370 CTTTGGCTCAATATGGACAGTGGAGAAATGTGAAACAAGTCAACACACACACACAAACC 429
Db |||||
101 LeuLeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThrAspThrGluThr 120
QY 430 GCCGTGTCAACGTCACATATGCAACAGAGAGAGCAAAATAGCCATCGAGAGACTA 489
Db |||||
121 AlaValValAsnValThrTyrAlaThrArgGluGluAlaLysIleAlaMetGluLysLeu 140
QY 490 AGCGGGCATCAGTTTGAGAACTACTCTCTCAAGATTTCTACATCCCGGATGAAGAGTGTG 549
Db |||||
141 SerGlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIleProAspGluGluVal 160
QY 550 AGCTCCCTTCGCCCTTCAGCGACCCAGCGCTGGGACCACTCTTCCCGGAGCAAGGC 609
Db |||||
161 SerSerProSerProGlnArgAlaGlnArgGlyAspHisSerSerArgGluGlnGly 180
QY 610 CACGCCCTTGGGGACATTCCTCAGGCGACAGACAGATGATGATTCCTCGGATCTCGTGC 669
Db |||||
181 HisAlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeuVal 200
QY 670 CCCACCCAGTTTGTGTGGTCCATCATCGAAAGAGGGCTTGACCATAAAGAACATCATCT 729
Db |||||
201 ProThrGlnPheValGlyAlaIleGlyLysGluGlyLeuThrIleLysAsnIleThr 220
QY 730 AAGCAGACCCAGTCCCGGTAGATATCCATAGAAAGAGAACTCTGGAGCTCCAGAGAG 789
Db |||||
221 LysGlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGlyAlaAlaGluLys 240
QY 790 CTTGTCCATTCATGCCACCCAGAGGGACTCTGAGCATGCCGATGATCTTGAA 849
Db |||||
241 ProValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeuGlu 260
QY 850 ATCATGCGAAGACGACGATGAGACCAACTAGCCGAGAGAGATTCCTCTGAAATCTTG 909
Db |||||
261 IleMetGlnLysGluAlaAspGluThrLysLeuAlaGluGluLeuLysIleLeu 280
QY 910 GCACCAATGGCTTGGTGGAAAGACTGATGAAAGAGAGAGAGAGAAATTTGAAGAAAT 969
Db |||||
281 AlaHisAsnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLysIle 300
QY 970 GAACATGAAACAGGACCAAGATAACAATCTCATCTTTCAGGATTTGAGCATATACAAC 1029
Db |||||
301 GluHisGluThrGlyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsn 320
QY 1030 CCGGAAAGAACCCATCACTGTGAAGGGCACAGTTGAGGCCCTGTGCCAGTCTCAGATPAG 1089
Db |||||
321 ProGluArgThrIleThrValLysGlyThrValGluAlaCysAlaSerAlaGluIleGlu 340
QY 1090 ATTATGAAGAGCTGCTGAGGCTTTGAAATGATATGCTGGCTGTTAAC 1140
Db |||||
341 IleMetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaValAsn 357

RESULT 6
AAU16166
ID AAU16166 standard; protein; 171 AA.
XX
AC AAU16166;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human novel secreted protein, Seq ID 1119.

XX Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;
KW cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;
KW antibacterial; virucide; fungicide; ophthalmological; vulnery;
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; skin ageing; food additive;
XX preservative; antiproliferative.
OS Homo sapiens.
XX
XX WO200155322-A2.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001341.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216840P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225113P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225477P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 18-AUG-2000; 2000US-0225759P.
PR 22-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234977P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.

PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0251990P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-488783/53.
DR N-PSDB; AAS26153.
XX
XX New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives.
XX
XX Claim 11; SEQ ID NO 1119; 980pp; English.
XX
XX The invention relates to isolated nucleic acid molecules and their
CC encoded secreted proteins. The nucleic acids and proteins are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC alleviating symptoms associated with the disorders and in diagnostic
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays
CC (ELISA). Disorders which are diagnosed or treated include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. cardiac
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
CC and many other disorders listed in the specification. The polypeptides
CC can also be used to aid wound healing and epithelial cell proliferation,
CC to prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence represents a novel secreted protein of the invention. Note: The
CC sequence data for this patent did not form part of the printed

Alignment Scores:
Pred. No.: 1e-142 Length: 171
Score: 160.00 Matches: 160
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.29% Indels: 0
DB: 4 Gaps: 0

US-09-270-437D-6 (1-3412) x AAU16166 (1-171)

QY 1387 GCGCCATCATCGGAAGAGGGGCGACATCAACAGCTGCGAGATTCCCGGAGCC 1446
DB 12 GlyAlaIleIleGlyIleGlyIleGlyAlaHisIleLysGlnLeuAlaArgPheAlaGlyAla 31
QY 1447 TCTATCAAGATTGCCCTCGGAAGGCCGAGAGCTGAGAAAGGATGTCATCATCAC 1506
DB 32 SerIleLysIleAlaProlaGluGlyProAspValSerGluArgMetValIleIleThr 51
QY 1507 GGGCCACCGAAGCCAGTTCAGGCCCGGAGCCAGGATCTTTGGGAAATCGAAAGAGGAA 1566
DB 52 GlyProGluAlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluGlu 71
QY 1567 AACTCTTTAAACCCCAAGAGAGTGAAGTGGAGCGCATATCAGAGTCCCTCTCC 1626
DB 72 AsnPheAsnProLysGluGluValLysLeuGluAlaHisIleArgValProSerSer 91

QY 1627 ACAGCTGCGCGGTGATTGGCAAGGTGGCAAGCCGTGACGAACTGCAGAACTTAACC 1686
DB 92 ThrAlaGlyArgValIleIleGlyLysGlyLysThrValAsnGluLeuGlnAsnLeuThr 111
QY 1687 AGTGCAGAAAGTCATCGTGCCTCGTGACCAACGCCAGATGAGGAAATGAGGAAATGATC 1746
DB 112 SerAlaGluValIleValProArgAspGlnThrProAspGluAsnGluGluValIleVal 131
QY 1747 AGAATTATCGGGCACTTCTTTTGTAGCCAGACTGCACAGCCCAAGATCAGGGAATTTGTA 1806
DB 132 ArgIleIleGlyHisPheAlaSerGlnThrAlaGlnArgLysIleArgGluIleVal 151
QY 1907 CAACAGGTGAGCAGCAGGAGCAGCAATAACCTCAGGAGTGCCTCACAGCGAGCAAG 1866
DB 152 GlnGlnValLysGlnGlnGlnGlnLysTyrProGlnGlyValAlaSerGlnArgSerLys 171
RESULT 7
ABU55235
ID ABU55235 standard; protein; 171 AA.
XX
AC ABU55235;
XX
DT 18-MAR-2003 (first entry)
XX
DE Human novel polypeptide #322.
XX
KW Human; neural disorder; immune system disorder; renal disorder;
KW muscular disorder; respiratory disease; reproductive disorder;
KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
KW hyperproliferative disorder; inflammatory disease; allergic reaction;
KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
KW haemostatic; antiarteriosclerotic.
XX
OS Homo sapiens.
XX
PN US2002132753-A1.
XX
PD 19-SEP-2002.
XX
PF 17-JAN-2001; 2001US-00764864.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214886P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0217496P.
PR 26-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 08-SEP-2000; 2000US-0229513P.
PR 21-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.

```

PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX
PA (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
Rosen CA, Ruben SM, Barash SC;
XX
WPI; 2003-147444/14.
DR N-PSDB; ABX73494.
XX
New polypeptides and nucleic acids, useful in gene therapy for treating,
PT inhibiting or preventing e.g. neural, immune system, muscular,
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
PT renal disorders.
XX
Claim 11; SEQ ID NO 1119; 402pp; English.
XX
The invention relates to human novel polypeptides and their associated
CC polynucleotides. The polypeptides and polynucleotides are useful in gene
CC therapy for treating, inhibiting or preventing neural disorders, immune
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
CC infarction) and cancerous diseases. Sequences ABU54914-ABU55699 and
CC ABU55748 represent human novel polypeptides of the invention
XX
SQ Sequence 171 AA;

```

```

Alignment Scores:
Pred. No.: 1e-142 Length: 171
Score: 160.00 Matches: 160
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.29% Indels: 0
DB: 6 Gaps: 0

```

US-09-270-437D-6 (1-3412) x ABU5235 (1-171)

```

QY 1387 GGGCCATCATCGGGAAGGGGGGACACATCAACAGCTGCGAGATTCGCGGAGCC 1446
Db 12 GYAlaIleIleIleGlyLysGlyAlaHisIleGlyGlnLeuAlaArgPheAlaGlyAla 31
QY 1447 TCTATCAGATTCCTCGGGAAGGCCCGACAGCTGAGGAAAGGATGTCATCATCACCC 1506
Db 32 SerIleLysIleAlaProAlaGluGlyProAspValSerGluArgMetValIleIleThr 51

```

```

QY 1507 GGGCCACCGGAAGCCAGTTCAAGGCCCCAGGAGCGGATCTTTGGGAAACTGAAGAGGAA 1566
Db 52 GlyProGluAlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluGlu 71
QY 1567 AACTTCTTTAACCCCAAGGAAGTGAAGCTGGAAGCGCATATCAGAGTGCCCTCTTCC 1626
Db 72 AsnPhePheAsnProLysGluGluValLysLeuAlaHisIleArgValProSerSer 91
QY 1627 ACAGCTGGCCGGGTGATTGGCAAGGTGGCAAGACCGTGAACGAACCTGCAAGACTTAACC 1686
Db 92 ThrAlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsnLeuThr 111
QY 1687 AGTGCAGAAGTCATCGTGGCTCTGTCGACCAACCCAGATCAAAATGAGGAAGTGCATGTC 1746
Db 112 SerAlaGluValIleValProArgAspGlnThrProAspGluAsnGluGluValIleVal 131
QY 1747 AGAATTATCGGCACACTCTTTGCTAGCCAGACTGCACAGCGCAAGATCAGGGAAATTGTA 1806
Db 132 ArgIleIleGlyHisPhePheAlaSerGlnThrAlaGlnArgLysIleArgGluIleVal 151
QY 1807 CAACAGGTGAAGCAGCAGGAGCAGAAATACCTCAGGAGTCGCCTCAGACGCGCAGCAAG 1866
Db 152 GlnGlnValLysGlnGlnGluGlnLysTyrProGlnGlyValAlaSerGlnArgSerLys 171
XX
RESULT 8
AAG03261
ID AAG03261 standard; protein; 97 AA.
XX
AC AAG03261;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein, SEQ ID NO: 7342.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping.
XX
OS Homo sapiens.
XX
PN EP1033401-AA.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-00200610.
XX
PR 26-FEB-1999; 99US-0122487P.
XX
PA (GEST ) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
WPI; 2000-500381/45.
DR N-PSDB; AAC03267.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
Claim 13; SEQ ID NO 7342; 71pp + Sequence Listing; English.
XX
The present sequence is a polypeptide encoded by one of a large number of
5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs were
prepared from total human RNAs or polyA+ RNAs derived from 30 different
tissues. EST sequences usually correspond mainly to the 3' untranslated
region (UTR) of the mRNA because they are often obtained from oligo-dT
primed cDNA libraries. Such ESTs are not well suited for isolating cDNA
sequences derived from the 5' ends of mRNAs and even in those cases where
longer cDNA sequences have been obtained, the full 5' UTR is rarely
included. 5' ESTs are derived from mRNAs with intact 5' ends and can
therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs
are also used in diagnostic, forensic, gene therapy and chromosome
mapping procedures. They are used to obtain upstream regulatory sequences
and to design expression and secretion vectors

```

```
XX SQ Sequence 97 AA;
Alignment Scores:
Pred. No.: 1.02e-82 Length: 97
Score: 97.00 Matches: 97
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.66% Indels: 0
DB: 3 Gaps: 0

US-09-270-437D-6 (1-3412) x AAG03261 (1-97)
QY 70 ATGATGACACAGCTTACATCGGAACTGACCCCGCCGTCACCGCGACCTCCGG 129
Db 1 MetMetAsnLysLeuTyIleGlyAsnLeuSerProAlaValThrAlaAspLeuArg 20
QY 130 CAGCTCTTTGGGGCAGGAAGCTGCCCTGGCGGACAGAGTCCCTGCTGAAGTCCGGCTAC 189
Db 21 GlnLeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLysSerGlyTyr 40
QY 190 GCCTTCGTGGACTACCCGACACAGAACTGGGCGCATCCGCGCATCGAGACCTCTCCGGGT 249
Db 41 AlaPheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGluThrLeuSerGly 60
QY 250 AAAGTGGAAATTCAGTGGAAATCATGGAAGTTGATTACTCAGTCTCTAAAAGCTAAGG 309
Db 61 LysValGluLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLeuArg 80
QY 310 AGCAGGAAATTCAGATTCGAAACATCCCTCTCCACCTGCAGTGGAGGTG 360
Db 81 SerArgLysIleGlnIleArgAsnIleProHisLeuGlnTrpGluVal 97

RESULT 9
AAU16164
ID AAU16164 standard; protein; 192 AA.
AC AAU16164;
DT 07-NOV-2001 (first entry)
DE Human novel secreted protein, Seq ID 1117.
KW Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;
KW cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;
KW antibacterial; virucide; fungicide; optalmalogical; vulnerary;
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; skin ageing; food additive;
KW preservative; antiproliferative.
XX OS Homo sapiens.
XX PN WO20015322-A2.
XX XX
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US001341.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 08-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0232402P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
```


PR	07-JUL-2000;	2000US-0216880P.	CC	leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
PR	11-JUL-2000;	2000US-0217487P.	CC	appendicitis), allergic reactions and conditions (e.g. asthma), blood
PR	11-JUL-2000;	2000US-0217496P.	CC	related disorders (e.g. thrombosis, atherosclerosis and myocardial
PR	14-JUL-2000;	2000US-0218290P.	CC	infection) and cancerous diseases. Sequences ABU54914-ABU55699 and
PR	26-JUL-2000;	2000US-0220963P.	CC	ABU55748 represent human novel polypeptides of the invention
PR	26-JUL-2000;	2000US-0220964P.	XX	
PR	14-AUG-2000;	2000US-0224518P.	SQ	Sequence 192 AA;
PR	14-AUG-2000;	2000US-0224519P.		
PR	14-AUG-2000;	2000US-0225267P.		
PR	14-AUG-2000;	2000US-0225268P.		
PR	14-AUG-2000;	2000US-0225270P.		
PR	14-AUG-2000;	2000US-0225447P.		
PR	14-AUG-2000;	2000US-0225757P.		
PR	14-AUG-2000;	2000US-0225758P.		
PR	22-AUG-2000;	2000US-0226868P.		
PR	30-AUG-2000;	2000US-0228924P.		
PR	01-SEP-2000;	2000US-0229287P.		
PR	01-SEP-2000;	2000US-0229343P.		
PR	01-SEP-2000;	2000US-0229344P.		
PR	01-SEP-2000;	2000US-0229345P.		
PR	05-SEP-2000;	2000US-0229509P.		
PR	08-SEP-2000;	2000US-0229513P.		
PR	21-SEP-2000;	2000US-0231413P.		
PR	21-SEP-2000;	2000US-0234223P.		
PR	21-SEP-2000;	2000US-0234274P.		
PR	25-SEP-2000;	2000US-0234997P.		
PR	27-SEP-2000;	2000US-0235834P.		
PR	29-SEP-2000;	2000US-0236327P.		
PR	29-SEP-2000;	2000US-0236367P.		
PR	29-SEP-2000;	2000US-0236368P.		
PR	29-SEP-2000;	2000US-0236369P.		
PR	29-SEP-2000;	2000US-0236570P.		
PR	02-OCT-2000;	2000US-0236802P.		
PR	02-OCT-2000;	2000US-0237037P.		
PR	02-OCT-2000;	2000US-0237038P.		
PR	02-OCT-2000;	2000US-0237039P.		
PR	02-OCT-2000;	2000US-0237040P.		
PR	13-OCT-2000;	2000US-0239935P.		
PR	20-OCT-2000;	2000US-0240960P.		
PR	20-OCT-2000;	2000US-0241785P.		
PR	20-OCT-2000;	2000US-0241809P.		
PR	01-NOV-2000;	2000US-0244617P.		
PR	17-NOV-2000;	2000US-0249299P.		
PR	08-DEC-2000;	2000US-0251856P.		
PR	08-DEC-2000;	2000US-0251868P.		
PR	08-DEC-2000;	2000US-0251869P.		
XX	(ROSE/) ROSEN C A.			
PA	(RUBE/) RUBEN S M.			
PA	(BARA/) BARASH S C.			
XX	Rosen CA, Ruben SM, Barash SC;			
PI				
XX	WPI; 2003-147444/14.			
DR	N-PSDB; ABX73492.			
XX				
PT	New polypeptides and nucleic acids, useful in gene therapy for treating,			
PT	inhibiting or preventing e.g. neural, immune system, muscular,			
PT	respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or			
PT	renal disorders.			
XX				
XX	Claim 11; SEQ ID NO 1117; 402pp; English.			
XX				
CC	The invention relates to human novel polypeptides and their associated			
CC	polynucleotides. The polypeptides and polynucleotides are useful in gene			
CC	therapy for treating, inhibiting or preventing neural disorders, immune			
CC	system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis			
CC	and multiple sclerosis), muscular disorders, respiratory diseases (e.g.			
CC	nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,			
CC	gastrointestinal disorders, pulmonary disorders, cardiovascular disorders,			
CC	(e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left			
CC	heart syndrome), renal disorders (e.g. acute kidney failure and end-stage			
CC	renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and			

CC	leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
CC	appendicitis), allergic reactions and conditions (e.g. asthma), blood
CC	related disorders (e.g. thrombosis, atherosclerosis and myocardial
CC	infection) and cancerous diseases. Sequences ABU54914-ABU55699 and
CC	ABU55748 represent human novel polypeptides of the invention
XX	
SQ	Sequence 192 AA;
Alignment Scores:	
Pred. No.:	8.07e-62
Score:	75.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	6.70%
DB:	6
US-09-270-437D-6 (1-3412) x ABU55233 (1-192)	
QY	994 ACAATCTCATCTTTGAGGATTGAGCATATACACCGGAAAGAACCATCTGTGAAG 1053
Db	72 ThrileSerSerLeuGlnAspLeuSerileTyraSnProGluArgThrIleThrVallys 91
QY	1054 GGCACAGTTGAGGCTGTGCGCAGTGTGAGATAGATATCAAGAGCTGCGTGAGGCC 1113
Db	92 GlyThrValGluAlaCysAlaSerAlaGluIleGluIleMetLysLysLeuArgGluAla 111
QY	1114 TTTGAAATGATATGCTGCTGTTAAACCAACAGCAATCTGATCCCGGGTTGAACCTC 1173
Db	112 PheGluAsnAspMetLeuAlaValAsnGlnGlnAlaAsnLeuIleProGlyLeuAsnLeu 131
QY	1174 AGCGCACTTGGCATCTTTTCAACAGGACTGTCCGTGCTATCTCCA 1218
Db	132 SerAlaLeuGlyIlePheSerThrGlyLeuSerValLeuSerPro 146
RESULT 11	
AAU16583	
ID	AAU16583 standard; protein; 171 AA.
XX	
AC	AAU16583;
XX	
DT	07-NOV-2001 (first entry)
DE	
XX	Human novel secreted protein, Seq ID 1536.
KW	Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;
KW	cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;
KW	antibacterial; virucide; fungicide; ophthalmological; vulnerary;
KW	secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW	cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW	cerebral ischaemia; angiogenesis; nervous system disorder;
KW	Alzheimer's disease; infection; ocular disorder; corneal infection;
KW	wound healing; epithelial cell proliferation; skin ageing; food additive;
KW	preservative; antiproliferative.
XX	
OS	Homo sapiens.
XX	
PN	WO2001155322-A2.
XX	
PD	02-AUG-2001.
XX	
PF	17-JAN-2001; 2001WO-US001341.
XX	
PR	31-JAN-2000; 2000US-0179065P.
PR	04-FEB-2000; 2000US-0180628P.
PR	24-FEB-2000; 2000US-0184664P.
PR	02-MAR-2000; 2000US-0186350P.
PR	16-MAR-2000; 2000US-0189874P.
PR	17-MAR-2000; 2000US-0190076P.
PR	18-APR-2000; 2000US-0198123P.
PR	19-MAY-2000; 2000US-0205151P.
PR	07-JUN-2000; 2000US-0209467P.
PR	28-JUN-2000; 2000US-0214886P.
PR	30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225575P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 08-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.

PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-488783/53.

N-PSDB; AAS26570.

PT New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives.

XX Claim 11; SEQ ID NO 1536; 980pp; English.

XX The invention relates to isolated nucleic acid molecules and their
CC encoded secreted proteins. The nucleic acids and proteins are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. Antibodies to the proteins can also be used in
CC alleviating symptoms associated with the disorders and in diagnostic

CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbant assays
CC (ELISA). Disorders which are diagnosed or treated include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiodenesis,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
CC and many other disorders listed in the specification. The polypeptides
CC can also be used to aid wound healing and epithelial cell proliferation,
CC to prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence represents a novel secreted protein of the invention. Note: The
CC sequence data for this patent did not form part of the printed

Alignment Scores:
Pred. No.: 3.03e-53 Length: 171
Score: 66.00 Matches: 66
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.89% Indels: 0
DB: 4 Gaps: 0

US-09-270-437D-6 (1-3412) x AAUI6583 (1-171)

QY 1609 ATCAGAGTCCCTCTTCCACAGCTGCGGGTGATTGCGAAAGTGGCAAGCCGTGAAC 1668
Db 86 lIeARgValProSerSerThrAlaGlyArgValIleGlyLysGlyThrValAsn 105
QY 1669 GAATCGCAGAACTTAACAGTGCAGAGTATCGTCGCTCGTGCACAAACGCCAGATGAA 1728
Db 106 GluLeuGlnAsnLeuThrSerAlaGluValIleValProArgAspGlnThrProAspGlu 125
QY 1729 ATGAGGAGTATCGTCTCAGATTATCGGGCTCTTTGCTAGCCAGACTGCACAGCGC 1788
Db 126 AsnGluGluValIleValArgIleIleGlyHisPheAlaSerGlnThrAlaGlnArg 145
QY 1789 AAGATCAGGAAATTTGTA 1806
Db 146 LysIleArgGluIleVal 151

RESULT 12
ABU55652
ID ABU55652 standard; protein; 171 AA.

AC ABU55652;

DT 18-MAR-2003 (first entry)

XX Human novel polypeptide #739.

XX Human: neural disorder; immune system disorder; renal disorder;
KW muscular disorder; respiratory disease; reproductive disorder;
KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
KW hyperproliferative disorder; inflammatory disease; allergic reaction;
KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
KW haemostatic; antiarteriosclerotic.

XX Homo sapiens.

OS US2002132753-A1.

XX 19-SEP-2002.

XX 17-JAN-2001; 2001US-00764864.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 28-JUN-2000; 2000US-0214886P.

PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 17-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.

(ROSE/) ROSEN C A.

(RUBE/) RUBEN S M.

(BARA/) BARASH S C.

Rosen CA, Ruben SM, Barash SC;

WPI; 2003-147444/14.

N-PSDB; ABX73911.

XX New polypeptides and nucleic acids, useful in gene therapy for treating,
PA inhibiting or preventing e.g. neural, immune system, muscular,
PA respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
PT renal disorders.

XX Claim 11; SEQ ID NO 1536; 402pp; English.

XX The invention relates to human novel polypeptides and their associated
CC polynucleotides. The polypeptides and polynucleotides are useful in gene
CC therapy for treating, inhibiting or preventing neural disorders, immune
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage

CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
CC infarction) and cancerous diseases. Sequences ABU54914-ABU55699 and
CC ABU55748 represent human novel polypeptides of the invention
XX
SQ Sequence 171 AA;

Alignment Scores:
Pred. No.: 3.03e-53 Length: 171
Score: 66.00 Matches: 66
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.8% Indels: 0
DB: 6 Gaps: 0

US-09-270-437D-6 (1-3412) x ABU55652 (1-171)

QY 1609 ATCAGAGTGCCTCTTCCACAGCTGGCGGGTGTATGCGCAAGGTGGCAAGCGTGAAC 1668
DB 86 IleArgValProSerSerThrAlaGlyArgValIleGlyLysGlyLysThrValAsn 105
QY 1669 GAACGTGAGAACTTAACCAAGTCAGAAATCATCGTCTGTCGACCAAGCCAGATGAA 1728
DB 106 GluLeuGlnAsnLeuThrSerAlaGluValIleValProArgAspGlnThrProAspGlu 125
QY 1729 AATGAGCAAGTATCGTCAGAAATTCATCGGGCACTTCTTTGTACCCAGACTGCACAGCGC 1788
DB 126 AsnGluGluValIleValArgIleGlyHisPheAlaSerGlnThrAlaGlnArg 145
QY 1789 AAGATCAGGGAATTGTA 1806
DB 146 LysIleArgGluIleVal 151

RESULT 13
ABU89799
ID ABU89799 standard; protein; 555 AA.
XX
AC ABU89799;
DT 10-JUL-2003 (first entry)
XX
DE Novel human protein NOV14a.
XX
KW Human; cytostatic; DAPK3-Agonist; DAPK3-Antagonist; cancer; NOV.
XX
OS Homo sapiens.

XX
PN WO2003031571-A2.
XX
PD 17-APR-2003.
XX
PF 02-OCT-2002; 2002WO-US031357.
XX
PR 05-OCT-2001; 2001US-0327454P.
PR 09-OCT-2001; 2001US-0327917P.
PR 09-OCT-2001; 2001US-0328029P.
PR 09-OCT-2001; 2001US-0328056P.
PR 12-OCT-2001; 2001US-0328056P.
PR 15-OCT-2001; 2001US-0329414P.
PR 17-OCT-2001; 2001US-0330142P.
PR 22-OCT-2001; 2001US-0341058P.
PR 24-OCT-2001; 2001US-0343629P.
PR 29-OCT-2001; 2001US-0349375P.
PR 01-NOV-2001; 2001US-0346557P.
PR 25-JUN-2002; 2002US-0391342P.
PR 01-OCT-2002; 2002US-00262445.

XX
PA (CURA-) CURAGEN CORP.
XX
PI Alsobrook JP, Burgess CE, Catterton E, Chant JS, Chaudhuri A;
PI Edinger SR, Gerlach VL, Giot L, Gorman L, Guo X, Kekuda R;

PI Mezes PS, Millet I, Ooi CE, Patturajan M, Rieger DK, Spytek KA;
PI Taupier RJ, Zerhusen BD, Zhong H, Zhong M;
XX
DR WPI; 2003-381704/36.
DR N-PSDB; ACA90176.
XX
PT New DAPK3 polypeptide, useful for preparing a composition for treating or
PT preventing e.g., cancer.
XX
PS Claim 2; Page 129; 253pp; English.

XX
CC The invention describes an isolated polypeptide comprising any of 33 90-
CC 1273 amino acid sequences (I) given in the specification or its mature
CC form, a sequence that is at least 95 % identical to (I), or a sequence
CC comprising one or more conservative substitutions in the amino acid
CC sequence of (I). The polypeptide is useful for preparing a composition
CC for treating or preventing e.g. cancer. This is the amino acid sequence
CC of a novel human NOV protein
XX

SQ Sequence 555 AA;

Alignment Scores:
Pred. No.: 2.23e-32 Length: 555
Score: 44.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.93% Indels: 0
DB: 6 Gaps: 0

US-09-270-437D-6 (1-3412) x ABU89799 (1-555)

QY 325 ATTGGAACATCCCTCCTCACCTGCGAGGTGTGTGATGAGCTTTGGCTCAATAT 384
DB 86 IleArgAsnIleProHisLeuGlnTrpGluValLeuAspGlyLeuLeuAlaGlnTr 105
QY 385 GGGACAGTGGAGAAATGTGGAAACAGTCAACACACACAGAACCCGCGTTGTCAAGTC 444
DB 106 GlyThrValGluAsnValGluGlnValAsnThrAspThrGluThrAlaValAsnVal 125
QY 445 ACATATGCAACA 456
DB 126 ThrTrAlaThr 129

RESULT 14
AAO23971
ID AAO23971 standard; protein; 187 AA.
XX
AC AAO23971;
DT 06-NOV-2003 (first entry)
XX
DE Human IGF-II mRNA binding protein 3.
XX
KW Anorectic; antiinflammatory; cardiant; hypotensive; antidiabetic;
KW neuroprotective; pharmaceutical composition; body-weight regulation;
KW thermogenesis; metabolic; obesity; Syndrome X; insulin-resistance;
KW eating disorder; cachexia; diabetes mellitus; hypertension; gallstone;
KW pancreatic dysfunction; arteriosclerosis; coronary heart disease;
KW hypercholesterolaemia; dyslipidaemia; osteoarthritis; ROS defence;
KW reactive oxygen species; neurodegenerative; mitochondrial; gene therapy;
KW human; IGF-II mRNA binding protein 3.

XX
OS Homo sapiens.
XX
PN WO2003061681-A2.
XX
XX 31-JUL-2003.
XX
PF 24-JAN-2003; 2003WO-EP000738.
XX
PR 25-JAN-2002; 2002EP-00001806.
PR 14-FEB-2002; 2002EP-00003473.
PR 28-FEB-2002; 2002EP-00004687.

PR 25-APR-2002; 2002EP-00009475.
PR 18-JUN-2002; 2002EP-00013329.
PR 30-DEC-2002; 2002EP-00029081.
XX
XX
XX
XX (DEVE-) DEVELOPEN ENTWICKLUNGSBIOLOGISCHE FORSCH.
XX
XX
XX Steuernagel A, Molitor A, Eulenberg K, Broemer G;
XX
XX WPI; 2003-627418/59.
XX N-PSDB; AAL57525.
XX
XX New pharmaceutical composition, useful for the manufacture of an agent
PT for diagnosing, treating or preventing disorders related to body-weight
PT regulation and thermogenesis, e.g., metabolic diseases such as obesity.
XX
XX Claim 3; Fig 7C; 144pp; English.
XX
XX The invention relates to a novel pharmaceutical composition comprising a
CC nucleic acid molecule or polypeptide which is a human homologue of a
CC drosophila melanogaster polypeptide or polynucleotide. The composition of
CC the invention may be utilised during the diagnosis, study, prevention and
CC treatment of diseases related to body-weight regulation and thermogenesis
CC including metabolic disorders such as obesity, Syndrome X and insulin-
CC resistance syndrome and eating disorders e.g. cachexia, diabetes
CC mellitus, hypertension, pancreatic dysfunctions, arteriosclerosis,
CC coronary heart disease, hypercholesterolaemia, dyslipidaemia,
CC osteoarthritis and gallstones. Furthermore, disorders related to reactive
CC oxygen species (ROS) defence may be addressed by the invention including
CC neurodegenerative disorders or mitochondrial disorders. Finally, the
CC composition of the invention may be useful in gene therapy. The current
CC sequence is that of the human IGF-II mRNA binding protein 3 protein of
CC the invention
XX
XX SQ Sequence 187 AA;

Alignment Scores:
Pred. No.: 187
Score: 32.00 Length: 187
Matches: 32
Conservative: 0
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.86%
Indels: 0
Gaps: 0

US-09-270-437D-6 (1-3412) x AAO23971 (1-187)

QY 1510 CCACCGGAAGCCAGTTCAGGCCAGGACCGATCTTTGGAACTGAAGAGGAAC 1569
Db 97 ProProGluAlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluGluAsn 116
QY 1570 TTCTTTAACCCCAAGAGAGAGTGAAGCTGGAAGCG 1605
Db 117 PhePheAsnProLysGluGluValLysLeuGluAla 128

RESULT 15
AAU16581
ID AAU16581 standard; protein; 47 AA.
XX
XX AAU16581;
XX
XX 07-NOV-2001 (first entry)
XX
XX Human novel secreted protein, Seq ID 1534.
XX
XX Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;
KW cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;
KW antibacterial; virucide; fungicide; ophthalmological; vulnery;
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; skin ageing; food additive;
KW preservative; antiproliferative.
XX

OS Homo sapiens.
XX
XX WO200155322-A2.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001341.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 16-MAR-2000; 2000US-0186350P.
PR 02-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216800P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 18-AUG-2000; 2000US-0225759P.
PR 22-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229387P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 21-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: July 16, 2004, 11:16:29 ; Search time 33 Seconds
(without alignments)
10675.625 Million cell updates/sec

Title: US-09-270-437D-6
Perfect score: 1120
Sequence: 1 ggcagcgagagcgagga.....aaccttgaaatgtttattt 3412

Scoring table: OLIGO

Xgapop 60.0 , Xgapext 60.0
Fgapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 389414 seqs, 51625971 residues

Word size: 1

Total number of hits satisfying chosen parameters: 663654

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p-model -DEV=xlp

-Q=/cgn2_1/USTO_spool_P/US09270437/runat.16072004.113128.13999/app query.fasta.1.3591

-DB=issued Patents AA -QFMT=faстан -SUFFIX=oligo.ra1 -MINMATCH=0.1 -LOOPCL=0

-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS-human40.cdi

-LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL

-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000

-USER=US09270437 @CGN 1_1_46@runat.16072004.113128.13999 -NCPUS=6 -ICPU=3

-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG

-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

1: Issued Patents AA.*
2: /cgn2_6/ptodata/2/iaa/5A COMB.pcp.*
3: /cgn2_6/ptodata/2/iaa/5B COMB.pcp.*
4: /cgn2_6/ptodata/2/iaa/6A COMB.pcp.*
5: /cgn2_6/ptodata/2/iaa/6B COMB.pcp.*
6: /cgn2_6/ptodata/2/iaa/6C COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	1.7	48	US-09-261-855-20	Sequence 20, Appl
2	19	1.7	577	US-09-261-855-2	Sequence 2, Appl
3	19	1.7	579	US-09-643-597-176	Sequence 176, Appl
4	19	1.7	579	US-09-643-597-348	Sequence 348, Appl
5	19	1.7	579	US-09-480-884A-176	Sequence 176, App
6	19	1.7	579	US-09-542-615A-176	Sequence 176, App
7	19	1.7	579	US-09-542-615A-348	Sequence 348, App
8	19	1.7	579	US-09-606-421B-176	Sequence 176, App
9	19	1.7	579	US-09-606-421B-348	Sequence 348, App
10	18	1.6	48	US-09-261-855-24	Sequence 24, Appl
11	16	1.4	49	US-09-261-855-22	Sequence 22, Appl
12	14	1.2	49	US-09-261-855-18	Sequence 18, Appl

13	1.2	47	3	US-09-261-855-21	Sequence 21, Appl
14	1.2	47	3	US-09-261-855-23	Sequence 23, Appl
15	1.0	47	3	US-09-261-855-19	Sequence 19, Appl
16	0.8	30	2	US-08-723-306-25	Sequence 25, Appl
17	0.8	30	2	US-08-723-306-30	Sequence 30, Appl
18	0.8	30	5	PCT-US96-10041-25	Sequence 30, Appl
19	0.8	30	5	PCT-US96-10041-30	Sequence 30, Appl
20	0.8	47	3	US-09-261-855-17	Sequence 17, Appl
21	0.8	81	4	US-09-489-039A-8805	Sequence 8805, Ap
22	0.8	156	4	US-09-252-991A-21912	Sequence 21912, A
23	0.8	399	4	US-09-252-991A-31280	Sequence 31280, A
24	0.8	491	1	US-08-206-176-4	Sequence 4, Appl
25	0.8	603	4	US-09-543-681A-6975	Sequence 6975, Ap
26	0.8	657	4	US-09-252-991A-28001	Sequence 28001, A
27	0.8	863	4	US-09-252-991A-21831	Sequence 21831, A
28	0.7	11	3	US-09-261-855-11	Sequence 11, Appl
29	0.7	11	3	US-09-261-855-13	Sequence 13, Appl
30	0.7	13	3	US-08-602-999A-90	Sequence 90, Appl
31	0.7	13	4	US-09-500-124-90	Sequence 90, Appl
32	0.7	31	3	US-08-602-999A-58	Sequence 58, Appl
33	0.7	31	4	US-09-500-124-58	Sequence 58, Appl
34	0.7	32	4	US-08-278-865-58	Sequence 58, Appl
35	0.7	74	4	US-09-621-976-6160	Sequence 6160, Ap
36	0.7	117	4	US-09-134-001C-5606	Sequence 5606, Ap
37	0.7	136	4	US-09-252-991A-26248	Sequence 26248, A
38	0.7	171	2	US-08-560-098A-53	Sequence 53, Appl
39	0.7	175	4	US-09-252-991A-31050	Sequence 31050, A
40	0.7	204	4	US-09-489-039A-7570	Sequence 7570, Ap
41	0.7	214	4	US-09-489-847-169	Sequence 169, App
42	0.7	228	4	US-09-724-864-42	Sequence 42, Appl
43	0.7	258	4	US-09-252-991A-21821	Sequence 21821, A
44	0.7	258	4	US-09-252-991A-27264	Sequence 27264, A
45	0.7	260	4	US-09-540-236-2747	Sequence 2747, Ap

ALIGNMENTS

RESULT 1

US-09-261-855-20
; Sequence 20, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; FILE OF INVENTION: (CRD-BP)
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 20
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-261-855-20

Alignment Scores:

Pred. No.: 7.89e-09
Score: 19.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.70%
DB: 3
Length: 48
Matches: 19
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0

US-09-270-437D-6 (1-3412) x US-09-261-855-20 (1-48)

QY 1630 GCTGCCCGCGTATTGGCAAGGTCGCAAGCCGTCGAACTGCAACTTAACC 1686

Db 8 AlaGlyArgValIleGlyGlyGlyThrValAsnGluLeuGlnAsnLeuThr 26

RESULT 2

US-09-261-855-2
; Sequence 2, Application US/09261855A

```

; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261.855A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 577
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-261-855-2

Alignment Scores:
Pred. No.:          5,25e-09          Length: 577
Score:              19.00             Matches: 19
Percent Similarity: 100.00%           Conservative: 0
Best Local Similarity: 100.00%        Mismatches: 0
Query Match:        1.70%             Indels: 0
DB:                  3                Gaps: 0

US-09-270-437D-6 (1-3412) x US-09-261-855-2 (1-577)

QY 1630 GCTGCCGGGTGATTGGCAAGGTGGCAAGCGGTGAACGAACTGACGAACCTTAACC 1686
Db 499 AlaGlyArgValIleGlyGlyGlyGlyThrValAsnGluLeuGlnAsnLeuThr 517

RESULT 3
US-09-643-597-176
; Sequence 176, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 176
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-643-597-176

Alignment Scores:
Pred. No.:          5,25e-09          Length: 579
Score:              19.00             Matches: 19
Percent Similarity: 100.00%           Conservative: 0
Best Local Similarity: 100.00%        Mismatches: 0
Query Match:        1.70%             Indels: 0
DB:                  4                Gaps: 0

US-09-270-437D-6 (1-3412) x US-09-643-597-176 (1-579)

QY 1489 AGATGGTCATCATCCCGGCCACCGGACCGGAGCCAGTTCAGGCCCGGACGGATC 1545
Db 452 ArgMetValIleIleThrGlyProGluAlaGlnPhelYsAlaGlnGlyArgIle 470

RESULT 4
US-09-643-597-348
; Sequence 348, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 348
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-643-597-348

Alignment Scores:
Pred. No.:          5,25e-09          Length: 579
Score:              19.00             Matches: 19
Percent Similarity: 100.00%           Conservative: 0
Best Local Similarity: 100.00%        Mismatches: 0
Query Match:        1.70%             Indels: 0
DB:                  4                Gaps: 0

US-09-270-437D-6 (1-3412) x US-09-643-597-348 (1-579)

QY 1489 AGATGGTCATCATCCCGGCCACCGGACCGGAGCCAGTTCAGGCCCGGACGGATC 1545
Db 452 ArgMetValIleIleThrGlyProGluAlaGlnPhelYsAlaGlnGlyArgIle 470

RESULT 5
US-09-480-884A-176
; Sequence 176, Application US/09480884A
; Patent No. 6482597
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; FILE REFERENCE: 210121.455C6
; CURRENT APPLICATION NUMBER: US/09/480,884A
; CURRENT FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 330
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 176
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-480-884A-176

Alignment Scores:
Pred. No.:          5,25e-09          Length: 579
Score:              19.00             Matches: 19
Percent Similarity: 100.00%           Conservative: 0
Best Local Similarity: 100.00%        Mismatches: 0
Query Match:        1.70%             Indels: 0
DB:                  4                Gaps: 0
```

US-09-270-437D-6 (1-3412) x US-09-480-884A-176 (1-579)

QY 1489 AGGATGGTCATCATCACCGGGCCACCGGAAGCCAGTTCAGGCCCGAGGACGGATC 1545
|||||
Db 452 ArgMetValIleIleThrGlyProGluAlaGlnPheLysAlaGlnGlyArgIle 470

RESULT 6

US-09-542-615A-176
; Sequence 176, Application US/09542615A
; Patent No. 6518256

GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; FILE REFERENCE: 210121.455C8
; CURRENT APPLICATION NUMBER: US/09/542, 615A
; CURRENT FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 176
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens

US-09-542-615A-176

Alignment Scores:

Pred. No.: 5.25e-09 Length: 579
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.70% Indels: 0
Gaps: 4

US-09-270-437D-6 (1-3412) x US-09-542-615A-176 (1-579)

QY 1489 AGGATGGTCATCATCACCGGGCCACCGGAAGCCAGTTCAGGCCCGAGGACGGATC 1545
|||||
Db 452 ArgMetValIleIleThrGlyProGluAlaGlnPheLysAlaGlnGlyArgIle 470

RESULT 7

US-09-542-615A-348

; Sequence 348, Application US/09542615A
; Patent No. 6518256

GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; FILE REFERENCE: 210121.455C8
; CURRENT APPLICATION NUMBER: US/09/542, 615A
; CURRENT FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 348
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens

US-09-542-615A-348

Alignment Scores:

Pred. No.: 5.25e-09 Length: 579
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.70% Indels: 0
Gaps: 4

US-09-270-437D-6 (1-3412) x US-09-542-615A-348 (1-579)

QY 1489 AGGATGGTCATCATCACCGGGCCACCGGAAGCCAGTTCAGGCCCGAGGACGGATC 1545
|||||
Db 452 ArgMetValIleIleThrGlyProGluAlaGlnPheLysAlaGlnGlyArgIle 470

RESULT 8

US-09-606-421B-176

; Sequence 176, Application US/09606421B
; Patent No. 6531315

GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606, 421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 176
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens

US-09-606-421B-176

Alignment Scores:

Pred. No.: 5.25e-09 Length: 579
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.70% Indels: 0
Gaps: 4

US-09-270-437D-6 (1-3412) x US-09-606-421B-176 (1-579)

QY 1489 AGGATGGTCATCATCACCGGGCCACCGGAAGCCAGTTCAGGCCCGAGGACGGATC 1545
|||||
Db 452 ArgMetValIleIleThrGlyProGluAlaGlnPheLysAlaGlnGlyArgIle 470

RESULT 9

US-09-606-421B-348

; Sequence 348, Application US/09606421B
; Patent No. 6531315

GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606, 421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 348

Alignment Scores:

Pred. No.: 5.25e-09 Length: 579
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0

```

; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-606-421B-348

Alignment Scores:
Pred. No.: 5,25e-09 Length: 579
Score: 19,00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.70% Indels: 0
DB: 4 Gaps: 0

US-09-270-437D-6 (1-3412) x US-09-606-421B-348 (1-579)

QY 1489 AGGATGTCATCATCACCGGGCCACCGGAAGCCCGAGTTCACGAGG
|||||
Db 452 ArgMetValIleIleThrGlyProGluAlaGlnPheLysA

RESULT 10
US-09-261-855-24
; Sequence 24, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINAN
; FILE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQU
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-261-855-24

Alignment Scores:
Pred. No.: 6.7e-08 Length: 48
Score: 18,00 Matches: 18
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.61% Indels: 0
DB: 3 Gaps: 0

US-09-270-437D-6 (1-3412) x US-09-261-855-24 (1-48)

QY 1630 GCTGCGCGGTGATTCGGCAAGGTGCGAAGACCGTGAACGACGAC
|||||
Db 8 AlaGlyArgValIleGlyGlyGlyThrValAsnGluL

RESULT 11
US-09-261-855-22
; Sequence 22, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINAN
; FILE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQU
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 49
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-261-855-22

Alignment Scores:
Pred. No.: 4.81e-06 Length: 49

```


Job time : 49 secs

```
QY      664 CTGGTCCGACCCAGTTTGTGTGTCGCATCATCGGAAG 702
Db      2 LeuValProThrGlnPheValGlyAlaIleIleGlyLys 14

RESULT 14
US-09-261-855-23
; Sequence 23, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-261-855-23

Alignment Scores:
Pred. No.:      0.00296      Length:      47
Score:          13.00      Matches:      13
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      1.16%      Indels: 0
DB:              3      Gaps: 0

US-09-270-437D-6 (1-3412) x US-09-261-855-23 (1-47)

QY      1432 AGATTCCGCGAGCCCTCTATCAAGATTGCCCTGCGGAA 1470
Db      24 ArgPheAlaGlyAlaSerIleIleAlaProAlaGlu 36

RESULT 15
US-09-261-855-19
; Sequence 19, Application US/09261855A
; Patent No. 6255055
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/261,855A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-261-855-19

Alignment Scores:
Pred. No.:      0.214      Length:      47
Score:          11.00      Matches:      11
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      0.98%      Indels: 0
DB:              3      Gaps: 0

US-09-270-437D-6 (1-3412) x US-09-261-855-19 (1-47)

QY      1378 CAGGCTGTGGCGCCCATCATCGGAGAGAGGGG 1410
Db      6 GlnAlaValGlyAlaIleIleGlyLysGly 16
```

Search completed: July 16, 2004, 11:30:45

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: July 16, 2004, 11:29:38 ; Search time 162.5 Seconds
(without alignments)

13125.623 Million cell updates/sec

Title: US-09-270-437D-6

Perfect score: 1120

Sequence: 1 ggcagcgaggagcgagga.....aacctgaaaatgtttattt 3412

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1285345 seqs, 312560633 residues

Word size: 1

Total number of hits satisfying chosen parameters: 2468542

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp

-Q=/cgn2_1/USPTO_spool_p/US09270437/runat_16072004_113130_14081/app_query.fasta_1.3591

-DB=Published Applications AA -QFMT=fastan -SUFFIX=oligo.rapp -MINMATCH=0.1

-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo

-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=1

-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0

-MAXLEN=2000000000 -USER=US09270437 @CGN 1 1 250 @runat_16072004_113130_14081

-NCPU=6 -ICPU=3 -NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100

-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60

-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

1	587	52.4	587	15	US-10-313-986-501	Sequence 501, App
2	511	45.6	620	9	US-09-764-864-1116	Sequence 1116, App
3	357	31.9	556	14	US-10-097-340-147	Sequence 147, App
4	357	31.9	556	16	US-10-648-593-182	Sequence 182, App
5	160	14.3	171	9	US-09-764-864-1119	Sequence 1119, App
6	75	6.7	192	9	US-09-764-864-1117	Sequence 1117, App
7	66	5.9	171	9	US-09-764-864-1536	Sequence 1536, App
8	44	3.9	555	15	US-10-262-445-40	Sequence 40, Appl
9	20	1.8	20	15	US-10-313-986-499	Sequence 499, App
10	20	1.8	47	9	US-09-764-864-1534	Sequence 1534, App
11	19	1.7	48	9	US-09-873-637-20	Sequence 20, Appl
12	19	1.7	81	14	US-10-117-982-476	Sequence 476, App
13	19	1.7	81	15	US-10-313-986-476	Sequence 476, App
14	19	1.7	250	9	US-09-764-864-1532	Sequence 1532, App
15	19	1.7	261	9	US-09-764-864-1114	Sequence 1114, App
16	19	1.7	422	16	US-10-408-765A-2088	Sequence 2088, App
17	19	1.7	577	9	US-09-873-637-2	Sequence 2, Appl
18	19	1.7	577	15	US-10-313-986-500	Sequence 500, App
19	19	1.7	579	9	US-09-735-705-176	Sequence 176, App
20	19	1.7	579	9	US-09-735-705-348	Sequence 348, App
21	19	1.7	579	9	US-09-850-716A-176	Sequence 176, App
22	19	1.7	579	9	US-09-850-716A-348	Sequence 348, App
23	19	1.7	579	9	US-09-897-778-176	Sequence 176, App
24	19	1.7	579	9	US-09-897-778-348	Sequence 348, App
25	19	1.7	579	9	US-09-897-778-446	Sequence 446, App
26	19	1.7	579	9	US-09-897-778-449	Sequence 449, App
27	19	1.7	579	10	US-09-466-396A-176	Sequence 176, App
28	19	1.7	579	12	US-10-007-700-176	Sequence 176, App
29	19	1.7	579	12	US-10-007-700-348	Sequence 348, App
30	19	1.7	579	12	US-10-007-700-446	Sequence 446, App
31	19	1.7	579	12	US-10-007-700-449	Sequence 449, App
32	19	1.7	579	14	US-10-117-982-176	Sequence 176, App
33	19	1.7	579	14	US-10-117-982-348	Sequence 348, App
34	19	1.7	579	14	US-10-117-982-446	Sequence 446, App
35	19	1.7	579	14	US-10-117-982-449	Sequence 449, App
36	19	1.7	579	14	US-10-117-982-480	Sequence 480, App
37	19	1.7	579	14	US-10-117-982-484	Sequence 484, App
38	19	1.7	579	15	US-10-313-986-176	Sequence 176, App
39	19	1.7	579	15	US-10-313-986-348	Sequence 348, App
40	19	1.7	579	15	US-10-313-986-446	Sequence 446, App
41	19	1.7	579	15	US-10-313-986-449	Sequence 449, App
42	19	1.7	579	15	US-10-313-986-480	Sequence 480, App
43	19	1.7	579	15	US-10-313-986-484	Sequence 484, App
44	19	1.7	586	9	US-09-850-716A-427	Sequence 427, App
45	19	1.7	586	9	US-09-897-778-427	Sequence 427, App

ALIGNMENTS

RESULT 1
US-10-313-986-501
; Sequence 501, Application US/10313986
; Publication No. US20030236209A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: McNabb, Andria
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tonglong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; CURRENT APPLICATION NUMBER: US/10/313,986
; FILE REFERENCE: 21021.455C19
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 501
; TYPE: PRT
; LENGTH: 587
; ORGANISM: Homo sapiens
US-10-313-986-501
Alignment Scores:

```
Pred. No.: 0 Length: 587
Score: 587.00 Matches: 587
Percent Similarity: 100.00% Conservative: 0
Best local Similarity: 100.00% Mismatches: 0
Query Match: 52.41% Indels: 0
DB: 15 Gaps: 0

US-09-270-437D-6 (1-3412) x US-10-313-986-501 (1-587)
QY 73 ATGAACAAGCTTTACATCGGGAACCTGAGCCCGCCGCTCACCGCCGACGACTCCGGCAG 132
Db 1 MetAsnLysLeuTyrIleGlyAsnLeuSerProAlaValThrAlaAspAspLeuArgGln 20
QY 133 CTCCTTTGGGACAGAAAGCTGCTCGGGGAGAGCTCTCTGAAGTCCGGTACGCC 192
Db 21 LeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLeuLysSerGlyTyrAla 40
QY 193 TTCGTGGACTACCCCGACAGAACTGGGCGCATCCGGCGCATCGAGACCTCTCCGGTAAA 252
Db 41 PheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGluThrLeuSerGlyLys 60
QY 253 GTGGAAATTCATGGGAAATCATGGAAGTTGATTACTCAGTCTCTAAAGCTAAGGACC 312
Db 61 ValGluLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLysLeuArgSer 80
QY 313 AGGAAATTCAGATTCGAAACATCCCTCTCACCTGCAGTGGGAGGTGTGGATGACTT 372
Db 81 ArgLysIleGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspGlyLeu 100
QY 373 TTGGCTCAATATGGGACAGTGGAGATGTGGAACAAGTCAACACAGACACAGAAACCGCC 432
Db 101 LeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThrAspThrGluThrAla 120
QY 433 GTTGTCACCTCACATATGCAACAAGAGAGCAAGCAAAATAGCCATGGAGAAGCTTAGC 492
Db 121 ValValAsnValTyrTyrAlaThrArgGluGluAlaLysIleAlaMetGluLysLeuSer 140
QY 493 GGGCATCAGTTTGAGACTACTCTTCAAGATTCTCTACATCCCGGATCAAGAGTGTAGC 552
Db 141 GlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIleProAspGluValSer 160
QY 553 TCCCTTCGCCCTCAGAGCCCGGAGTGGGACCACTCTTCCCGGAGCAAGCCAC 612
Db 161 SerProSerProGlnArgAlaGlnArgGlyAspHisSerSerArgGluGlnGlyHis 180
QY 613 GCCCTGGGGGACCTTCTCAGCCGACAGATGATTTCCTCCGCTCGGATCTCTGTCGCC 672
Db 181 AlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeuValPro 200
QY 673 AGCCAGTTTGTGGTGCATCATCGAAAGGAGGCTTCACCATAAAGAACATCACTAAG 732
Db 201 ThrGlnPheValGlyAlaIleGlyLysGluGlyLeuThrIleLysAsnIleThrLys 220
QY 733 CAGACCCAGTCCCGGTAGATATCCATAGAAAGAGAACTCTGGAGCTCGAGAGGCTT 792
Db 221 GlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGlyAlaAlaGluLysPro 240
QY 793 GTCCACCATCCATCCACCCAGAGGGGACTTCTGAGCATGCGGATGATCTTGAATC 852
Db 241 ValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeuGluIle 260
QY 853 ATGCAAGAGAGGCATGAGCAACCACTAGCCGGAAGAGATTCCTCTGAAAATCTTGCA 912
Db 261 MetGlnLysGluAlaAspGluThrLysLeuAlaGluIleProLeuLysIleLeuAla 280
QY 913 CACAATGCTGTTGGAAGACTGATTCGAAAGAGAGGAGCAAAATTTGAAGAAATGAA 972
Db 281 HisAsnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLysIleGlu 300
QY 973 CATGAACAGGACCAAGATAACAATCTCATCTTTTCAGGATTTGAGCATATACACCG 1032
Db 301 HisGluThrGlyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsnPro 320
```

```
QY 1033 GAAAGAACCATCACTGTGAAGGCACAGTTGAGGCGCTGTGCAGTGTGAGATGAGATT 1092
Db 321 GluArgThrIleThrValLysGlyThrValGluAlaCysAlaSerAlaGluIleGluIle 340
QY 1093 ATGAAGAGCTGCGTGAGGCTTTGAAAATCATATGCTGGCTGTTAAACAACAGCCAA 1152
Db 341 MetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaValAsnGlnAlaAsn 360
QY 1153 CTGATCCAGAGGTTGAACCTCAGCGCACTTGGCATCTTTCAACAGGACTGCTCGTGCTA 1212
Db 361 LeuIleProGlyLeuAsnLeuSerAlaLeuGlyIlePheSerThrGlyLeuSerValLeu 380
QY 1213 TCTCACAGCAGAGGCGCCGCGAGCTCCCGCGCTGCCCTACACCCCTTCACTACC 1272
Db 381 SerProProAlaGlyProArgGlyAlaProProAlaAlaProTyrHisProPheThr 400
QY 1273 CACTCCGATATCTCTCAGCTGTACCCCATCACAGTTCGCGAGATTTCGCGAGCTCTATC 1332
Db 401 HisSerGlyTyrPheSerSerLeuTyrProHisGlnPheGlyProPheProHisHis 420
QY 1333 CACTCTTATCCAGAGCAGGAGATTGTGAATCTCTTCATCCCAACCCAGCGTGTGGCGCC 1392
Db 421 HisSerTyrProGluGlnGluIleValAsnLeuPheIleProThrGlnAlaValGlyAla 440
QY 1393 ATCATCGGGAAGAGGGGCGCACACATCAACAGCTGCGGAGATTTCGCGAGCTCTATC 1452
Db 441 IleIleGlyLysLysGlyAlaHisIleLysGlnLeuAlaArgPheAlaGlyAlaSerIle 460
QY 1453 AAGATTGCCCTGCGGAAGGCGCAGAGCTGCGGAAAGATGCTCATCATCCCGGCGCA 1512
Db 461 LysIleAlaProAlaGluGlyProAspValSerGluArgMetValIleIleThrGlyPro 480
QY 1513 CCGGAAGCCCGCTCAAGCGCCAGGAGCTTTTGGGAACTCTTGGAACTGAAAGAGAACTTC 1572
Db 481 ProGluAlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluGluAsnPhe 500
QY 1573 TTTAAACCCCAAGAAAGTAGTGAAGCTGGAAGCGCATATCAGAGTGCCTCTTCCACAG 1632
Db 501 PheAsnProLysGluGluValLysLeuGluAlaHisIleArgValProSerSerThrAla 520
QY 1633 GCGCGGCTGATTGGCAAGGTTGGAAGCGGTGAAGCGTGAAGCAACTGACAACTTAACCA 1692
Db 521 GlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsnLeuThrSerAla 540
QY 1693 GAAAGTCATCGTCCCTCGTGACCAAGCGCAGATGAAATGAGAAAGTGCATCTCAGAAT 1752
Db 541 GluValIleValProArgAspGlnThrProAspGluAsnGluGluValIleValArgIle 560
QY 1753 ATCGGCGCACTTCTTTGTAGCCAGACTGCACAGCGCAAGATCAGGGAATTTGTAACA 1812
Db 561 IleGlyHisPhePheAlaSerGlnThrAlaGlnArgLysIleArgGluIleValGlnGln 580
QY 1813 GTGAAGCAGCAGGACAGAAA 1833
Db 581 ValLysGlnGlnGluGlnLys 587

RESULT 2
US-09-764-864-1116
; Sequence 1116, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1116
; LENGTH: 620
; TYPE: PRI
; ORGANISM: Homo sapiens
```

```
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (533)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1116

Alignment Scores:
Pred. No.: 0          Length: 620
Score: 511.00        Matches: 511
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 45.62%      Indels: 0
DB: 9                  Gaps: 0

US-09-270-437D-6 (1-3412) x US-09-764-864-1116 (1-620)

QY 70 ATGATCAACAGCTTTTACATCGGGAACCTGAGCCCGCGCTGACCCCGACGACCTCCGG 129
Db 22 MetMetAsnLysLeuTyrIleGlyAsnLeuSerProAlaValThrAlaAspLeuArg 41
QY 130 CAGCTCTTTGGGACAGGAAGTCCCGCTGGCGGGACAGGTCTCTGCTGAAGTCCGGCTAC 189
Db 42 GlnLeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLeuLysSerGlyTyr 61
QY 190 GCCTTCGTGGACTACCCCGACAGAACTGGGCCATCGCGGCCATCGAGACCTCTCGGCT 249
Db 62 AlaPheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGluThrLeuSerGly 81
QY 250 AAAGTGGAAATTCATCGGAAATCATGGAAATGATTACTCAGTCTCTAAAGAGTAAGG 309
Db 82 LysValGluLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLysLeuArg 101
QY 310 AGCAGAAATTCAGATTGAAACATCCCTCTCAGCTCGAGTGGAGGTGTTGGATGA 369
Db 102 SerArgLysIleGlnIleArgAsnIleProProHisLeuGlnTrpGluValLeuAspGly 121
QY 370 CTTTGGCTCAATATGGACAGTGGAGATGTGGAACAAGTCAACACAGACAGAAACC 429
Db 122 LeuLeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThrAspThrGluThr 141
QY 430 GCGTGTGTCAAGCTTCATATGCAACAGAGAGAGAGCAAAATAGCCATGGAGAAGCTA 489
Db 142 AlaValValAsnValThrTyrAlaThrArgGluGluAlaLysIleAlaMetGluLysLeu 161
QY 490 AGCGGCATCAGTTGAGAACTACTCTTCAAGATTCTCTACATCCCGGATGAAGAGTG 549
Db 162 SerGlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIleProAspGluVal 181
QY 550 AGCTCCCTTCGCGCCCTCAGCGAGCCCGCTGGGACCCACTCTTCCCGGAGCAAGC 609
Db 182 SerSerProSerProProGlnArgAlaGlnArgGlyAspHisSerSerArgGluGlnGly 201
QY 610 CAGCCCTTCGGGGCCACTCTCAGGCGCAGACAGATTGATTCCCGCTGCGGATCCCTGGTC 669
Db 202 HisAlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeuVal 221
QY 670 CCCACCCAGTTTGTGGTCCATCATCGGAAAGAGAGGCTTGACATAAGAACATCACT 729
Db 222 ProThrGlnPheValGlyAlaIleIleGlyLysGluGlyLeuThrIleLysAsnIleThr 241
QY 730 AAGCAGNCCAGTCCCGGCTAGATATCCATAGAAAGAGAACTCTGGAGCTGCAGAGAAG 789
Db 242 LysGlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGlyAlaAlaGluLys 261
QY 790 CTTGTCCACCATCCATGCCACCCAGAGGGCACTTCTGAAGCATGCCCATGATTCTTGAA 849
Db 262 ProValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeuGlu 281
QY 850 ATCATGAGAAAGAGGAGATGAGACCAAACTAGCCGAGAGAGATTCTCTGAAATCTTG 909
Db 282 IleMetGlnLysGluAlaAspGluThrLysLeuAlaGluIleProLeuLysIleLeu 301
QY 910 GCACACATCGCTTGGTGGAAAGACTGATTGGAAAGAGAGGAGAAATTTGAAGAAATT 969
```

```
Db 302 AlaHisAsnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLysLysIle 321
QY 970 GAACATGAACAGGAGGACCAAGATAACAATCTCATCTTTTCAGGATTTGAGCATATACAAC 1029
Db 322 GluHisGluThrGlyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsn 341
QY 1030 CCGGAAAGAACCATCATCTGTGAAGGCGACAGTTGAGCCCTGTGCCAGTGTCTGAGATAGAG 1089
Db 342 ProGluArgThrIleThrValLysGlyThrValGluAlaCysAlaSerAlaGluIleGlu 361
QY 1090 ATTATCAAGAGCTGCTGAGCCCTTTGAAATGATATCTGCTCTTAAACCAACAGGCC 1149
Db 362 IleMetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaValAsnGlnAla 381
QY 1150 AATCTCATCCAGGGTTGAACCTCAGCGACCTTGGCATCTTTTCAACAGGACTGTCGGTG 1209
Db 382 AsnLeuIleProGlyLeuAsnLeuSerAlaLeuGlyIlePheSerThrGlyLeuSerVal 401
QY 1210 CTATCTCCACGACGAGGCGCGCGGAGCTCCCGCTGCGCCCTTACCAACCCCTTCACT 1269
Db 402 LeuSerProProAlaGlyProArgGlyAlaProProAlaAlaProTyrHisProPheThr 421
QY 1270 ACCACTCCGGATCTCTCCAGCCCTTACCCCATCACAGTTTGGCCGCTCCCGCAT 1329
Db 422 ThrHisSerGlyTyrPheSerSerLeuTyrProHisGlnPheGlyProPheHis 441
QY 1330 CATCACTCTTATCCAGCAGGAGATTGTGAATCTCTTATCCCAACCCAGGCTGTGGGC 1389
Db 442 HisHisSerTyrProGluGlnGluIleValAsnLeuPheIleProThrGlnAlaValGly 461
QY 1390 GCATCATCGGAAGAGGGGCGACACATCAACAGCTGGCGAGATTCGCGAGGCTCT 1449
Db 462 AlaIleIleGlyLysGlyAlaHisIleLysGlnLeuAlaArgPheAlaGlyAlaSer 481
QY 1450 ATCAAGATTGCCCTCGGAGGCCCGACAGCTCAGGAGGAGGATGCTCATCATCAGCGG 1509
Db 482 IleLysIleAlaProAlaGluGlyProAspValSerGluArgMetValIleIleThrGly 501
QY 1510 CCACCGGAAGCCAGTCTCAAGGCCAGGACCGATCTTTGGGAAACTGAAAGAGGAAAC 1569
Db 502 ProProGluAlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluGluAsn 521
QY 1570 TTCTTTAAACCCAAAGAAAGTGAAGCTGGAA 1602
Db 522 PhePheAsnProLysGluGluValLysLeuGlu 532

RESULT 3
US-10-097-340-147
; Sequence 147, Application US/10097340
; Publication No. US20030087250A1
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNANAVARAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISSEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumel ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; CURRENT FILING DATE: 2002-03-14
```

; PRIOR APPLICATION NUMBER: 60/276, 025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325, 149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276, 026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324, 967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311, 732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325, 102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323, 580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 147
; LENGTH: 556
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-097-340-147

Alignment Scores:
Pred. No.: 9.24e-315 Length: 556
Score: 357.00 Matches: 357
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 31.87% Indels: 0
DB: 14 Gaps: 0

US-09-270-437D-6 (1-3412) x US-10-097-340-147 (1-556)

QY 70 ATGATGAACAAGCTTACATCGGAACCTGAGCCCGCCGTCACCGCGAGCACTCCGG 129
Db 1 MetMetAsnLysLeuThrIleGlyAsnLeuSerProAlaValThrAlaAspLeuArg 20
QY 130 CAGCTCTTTGGGGACAGGAAGTCGCCCTGGCGGACAGGTCCTGCTCAAGTCGGGTAC 189
Db 21 GlnLeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLeuLysSerGlyTyr 40
QY 190 GCCTTCGTGGACTACCCGACAGAACTCGGGCCATCCCGCCGATCGAGACCTCTCGGGT 249
Db 41 AlaPheValAspTyrProAspGlnAsnTrpAlaIleArgAlaIleGluThrLeuSerGly 60
QY 250 AAAGTGAATTCATGCGGAAATCATGGAAGTTGATTACTCAGTCTCTAAAAGCTAAGG 309
Db 61 LysValGluLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLeuArg 80
QY 310 AGCAGGAAATTCAGATTCGAAACATCCCTCTCAGTCGAGTGGAGGTGTGGATGGA 369
Db 81 SerArgLysIleGlnIleArgAsnIleProHisLeuGlnTrpGluValLeuAspGly 100
QY 370 CTTTGGCTCAATATGGACAGTGGAGATGTGGNACAGTCAAGTCAACACAGACAGAACCC 429
Db 101 LeuLeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThrAspThrGluThr 120
QY 430 GCCCTTGTCAACGTCACATATGCAACAGAGAAGCAAAATAGCCATCGAGAAGCTA 489
Db 121 AlaValValAsnValThrTyrAlaThrArgGluGluAlaLysIleAlaMetGluLysLeu 140
QY 490 AGCGGGCATCAGTTTGAAACTACTCTTCAAGATTTCTCATATCCCGGATCAAGAGGTG 549
Db 141 SerGlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIleProAspGluGluVal 160
QY 550 AGCTCCCTTCGCCCTCAGCGACCCGAGCTGGGACCACTCTTCCCGGAGCAAGGC 609
Db 161 SerSerProSerProGlnHarGalaGlnArgGlyAspHisSerSerArgGluGlnGly 180
QY 610 CACGCCCTGGGGCACTTCTCAGGCCACAGATTGATTTCCCGCTCGGATCCTGGTC 669
Db 181 HisAlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeuVal 200
QY 670 CCCACCAAGTTTGTGGTGGCCATCATCGGAAAGGAGGCTTGACCATTAAGAACATCACT 729

Db 201 ProThrGlnPheValGlyAlaIleIleGlyLysGluGlyLeuThrIleLysAsnIleThr 220
QY 730 AAGCAGACCCAGTCCCGGTAGATATCCATAAAGAGAACTCTGGAGCTGCAGAGAAG 789
Db 221 LysGlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGlyAlaAlaGluLys 240
QY 790 CCTGTCAACCATCCATGCCACCCAGAGGGACCTTCTAGAGCATGCCGATATCTTGA 849
Db 241 ProValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeuGlu 260
QY 850 ATCATGCAAGAAAGGAGGATGACCAAACTAGCCGAGAGATTCTCTCTGAAAATCTTG 909
Db 261 IleMetGlnLysGluAlaAspGluThrLysLeuAlaGluGluIleProLeuLysIleLeu 280
QY 910 GCACACAATGGCTTGGTGAAGACTGATTGGAAGAGGACGAAATTTGAAGAAATTT 969
Db 281 AlaHisAsnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLysLysIle 300
QY 970 GAACATGAACAGGAGCCAGATAACATCTCATCTTTGAGGATTTGAGCATATACAC 1029
Db 301 GluHisGluThrGlyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsn 320
QY 1030 CCGGAAAGAACCATCACTGTGAAGGACACAGTTGAGGCTGTGCAGTCTGAGATAGAG 1089
Db 321 ProGluArgThrIleThrValLysGlyThrValGluAlaCysAlaSerAlaGluIleGlu 340
QY 1090 ATTATGAAGAGCTGCGTGAGGCTTTGAAAATGATATGCTGGCTGTTTAAAC 1140
Db 341 IleMetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaValAsn 357

RESULT 4

US-10-648-593-182
; Sequence 182, Application US/10648593
; Publication No. US20040106132A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: IDENTIFICATION OF GENES FOR PREDICTING ACTIVITY OF COMPOUNDS THAT
; INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE KINASES AND/OR
; TITLE OF INVENTION: PROTEIN TYROSINE KINASE PATHWAYS IN BREAST CELLS
; FILE REFERENCE: D0273 NP
; CURRENT APPLICATION NUMBER: US/10/648,593
; CURRENT FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: 60/406,385
; PRIOR FILING DATE: 2002-08-27
; NUMBER OF SEQ ID NOS: 557
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 182
; LENGTH: 556
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-648-593-182

Alignment Scores:
Pred. No.: 9.24e-315 Length: 556
Score: 357.00 Matches: 357
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 31.87% Indels: 0
DB: 16 Gaps: 0

US-09-270-437D-6 (1-3412) x US-10-648-593-182 (1-556)

QY 70 ATGATGAACAAGCTTACATCGGAACCTGAGCCCGCCGTCACCGCGAGCACTCCGG 129
Db 1 MetMetAsnLysLeuThrIleGlyAsnLeuSerProAlaValThrAlaAspLeuArg 20
QY 130 CAGCTCTTTGGGGACAGGAAGTCGCCCTGGCGGACAGGTCCTGCTGAACTCCGGTAC 189
Db 21 GlnLeuPheGlyAspArgLysLeuProLeuAlaGlyGlnValLeuLeuLysSerGlyTyr 40
QY 190 GCCTTCGTGGACTACCCGACAGAACTCGGGCCATCCCGCCATCGAGACCTCTCGGGT 249

```
Db 41 AlaPheValAspTyrProAspGlnAsnTrpAlaIleAlaIleGluThrIleuSerGly 60
QY 250 AAAGTGGATTCGATCGGAAATCATGGAATGATTACTAGTCTCTAAAGCTAAGG 309
Db 61 LysValGluLeuHisGlyLysIleMetGluValAspTyrSerValSerLysLysLeuArg 80
QY 310 AGCAGGAAATTCAGATTTCGAACATCCCTCTCCTACCTGCGAGTGGAGGTGTGGATGGA 369
Db 81 SerArgLysIleGlnIleArgAsnIleProProHisLeuGlnTrpGluValIleuAspGly 100
QY 370 CTTTGGCTCAATATGGACAGTGGAGATGTGGAACAAGTCAACACAGACACAGAAACC 429
Db 101 LeuLeuAlaGlnTyrGlyThrValGluAsnValGluGlnValAsnThrAspThrGluThr 120
QY 430 GCGCTGTTCACAGTCACATATGCACACAGACAGAGCAAAAATAGCCATGGAGAGCTA 489
Db 121 AlaValValAsnValThrTyrAlaThrArgGluGluAlaLysIleAlaMetGluLysLeu 140
QY 490 AGCGGGCATCAGTTTGGAGAACTACTCTCTCAAGATTTCTCTACATCCCGGATGAGAGGTG 549
Db 141 SerGlyHisGlnPheGluAsnTyrSerPheLysIleSerTyrIleProAspGluGluVal 160
QY 550 AGCTCCCTTCGCCCCCTACGAGCCCGAGCCGCTGGGACCACTCTTCCCGGGAGCAAGGC 609
Db 161 SerSerProSerProProGlnArgAlaGlnArgGlyAspHisSerSerArgGluGlnGly 180
QY 610 CACGCCCTGGGGCACTCTCAGGCCAGACAGATTGATTTCCGCTGCGGATCCTGGTC 669
Db 181 HisAlaProGlyGlyThrSerGlnAlaArgGlnIleAspPheProLeuArgIleLeuVal 200
QY 670 CCCACCCAGTTTGTGTGGTCCATCATCGGAAGAGGCTTGACCATPAAAGACATCACT 729
Db 201 ProThrGlnPheValGlyAlaIleIleGlyLysGluGlyLeuThrIleLysAsnIleThr 220
QY 730 AAGCAGACCCAGTCCCGGTAGATATCCATAGAAAGAGAACTCTGGAGCTGCAGAGAAG 789
Db 221 LysGlnThrGlnSerArgValAspIleHisArgLysGluAsnSerGlyAlaAlaGluLys 240
QY 790 CTTGTACCATCCATCCACCCAGAGGACTCTCTGAAGCATGCCCATGCTCTTTGAA 849
Db 241 ProValThrIleHisAlaThrProGluGlyThrSerGluAlaCysArgMetIleLeuGlu 260
QY 850 ATCATGAGAAAGAGGAGATGAGACCAAACTAGCCGAGAGATTCCTCTGAAATCTTG 909
Db 261 IleMetGlnLysGluAlaAspGluThrLysLeuAlaGluGluLeuProLeuLysIleLeu 280
QY 910 GCACAAATGGCTGTGGTGAAGCTGATTGGAAGAGAGCCAGAAATTTGAAGAAAAT 969
Db 281 AlaHisAsnGlyLeuValGlyArgLeuIleGlyLysGluGlyArgAsnLeuLysLysIle 300
QY 970 GAACATGAACAGGAGCCAGATAACAATCTCTCTTTCAGGATTTGACCATATCAAC 1029
Db 301 GluHisGluThrGlyThrLysIleThrIleSerSerLeuGlnAspLeuSerIleTyrAsn 320
QY 1030 CCGGAAGAACCATCACTGTGAAGGGCAGAGTGTGAGCCCTGTGCCAGTGTGTGAGATAG 1089
Db 321 ProGluArgThrIleThrValLysGlyThrValGluAlaCysAlaSerAlaGluIleGlu 340
QY 1090 ATTATCAAGAGCTGGTGGAGGCTTTGAAATGATATGATGCTGCTTAAAC 1140
Db 341 IleMetLysLysLeuArgGluAlaPheGluAsnAspMetLeuAlaValAsn 357
```

RESULT 5

```
US-09-764-864-1119
; Sequence 1119, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT23
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
```

```
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1119
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-864-1119

Alignment Scores:
Pred. No.: 8,27e-136 Length: 171
Score: 160.00 Matches: 160
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14,29% Indels: 0
DB: 9 Gaps: 0

US-09-270-437D-6 (1-3412) x US-09-764-864-1119 (1-171)
QY 1387 GCGCCCATCATCGGAAAGAGGGGCACACATCAACAGCTGGCGAGATTCGCCGAGCC 1446
Db 12 GlyAlaIleIleGlyLysLysGlyAlaHisIleLysGlnLeuAlaArgPheAlaGlyAla 31
QY 1447 TCTATCAAGATTGCCCTCGGAAAGGCCACAGCTCAGCGAAAGATGGTCATCATCACC 1506
Db 32 SerIleLysIleAlaProAlaGluGlyProAspValSerGluArgMetValIleIleThr 51
QY 1507 GGGCCACCGGAAGCCCAAGTTCAAGGCCCGAGGGAGGATCTTGGGAAACTGAAAGAGAA 1566
Db 52 GlyProGluAlaGlnPheLysAlaGlnGlyArgIlePheGlyLysLeuLysGluGlu 71
QY 1567 AACTTCTTTAACCCCAAGAGTGAAGCTGGAAGCGCATATCAGAGTGCCTCTTCC 1626
Db 72 AsnPheAsnProLysGluGluValLysLeuGluAlaHisIleArgValProSerSer 91
QY 1627 ACAGCTGGCCGGTGATTGGCAAGAGTGGCAAGACCGTGAACCACTGCAAGACTTAACC 1686
Db 92 ThrAlaGlyArgValIleGlyLysGlyLysThrValAsnGluLeuGlnAsnLeuThr 111
QY 1687 AGTGCAAGACTCATCGTCTGCTGACCAACCCAGATGAAATGAGGAAGTATCGTC 1746
Db 112 SerAlaGluValIleValProArgAspGlnThrProAspGluAsnGluGluIleVal 131
QY 1747 AGAATATCGGCCTCTTTGCTAGCAGACTGCACAGCGCAAGATCAGGAAATGTA 1806
Db 132 ArgIleIleGlyHisPheAlaSerGlnThrAlaGlnArgLysIleArgGluIleVal 151
QY 1807 CAACAGTGAACAGCAGCAGGAGCAAGATACCTCAGGAGTGCCTCAGCGCAGCAAG 1866
Db 152 GlnGlnValLysGlnGlnGluGlnLysTyrProGlnGlyValAlaSerGlnArgSerLys 171

RESULT 6
US-09-764-864-1117
; Sequence 1117, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT23
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1117
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (71)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (147)
```

```

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (160)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (164)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (165)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (168)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-764-864-1117

```

Alignment Scores:		
Pred. No.:	1.3e-58	Length:
Score:	75.00	Matches:
Percent Similarity:	100.00%	Conservative:
Best Local Similarity:	100.00%	Mismatches:
Query Match:	6.70%	Indels:
DR:	9	Gaps:
		0
		0
		192
		75

US-09-270-437D-6 (1-3412) x US-09-764-864-1117 (1-192)

994	Qy	ACAAATCTCACTCTTTGGAGGATTTCAGCATATACACCGGAAGAACCATCACTGTAAG	1055
72	Db	ThrIleSerSerLeuGlnAspLeuSerIleIleAsnProGluArgThrIleThrValIys	91
1054	Qy	GGCACAGTTCAGGCCTCTGCCAGGCTGAGATAGAGATTATGAAGAGCTGCGTGAGGCC	1113
92	Db	GlyThrValGluAlaCysAlaSerAlaGluIleGluIleMetIysLeuArgGluAla	111
1114	Qy	TTTGAAATATGATCTGGCTGGTGTAAACCAACAGCCCAATCTGATCCCGAGGGTTGAACCTC	1173
112	Db	PheGluAsnAspMetLeuAlaValAsnGlnGlnAlaAsnLeuIleProGlyLeuAsnLeu	131
1174	Qy	AGCGCACTTGGCATCTTTTCAACAGGACTGTCCGTGTATCTCCA	1218
132	Db	SerAlaLeuGluIlePheSerThrGlyLeuSerValLeuSerPro	146

```

RESULT 7
US-09-764-864-1536
; Sequence 1536, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ23
; CURRENT APPLICATION NUMBER: US/09/764, 864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1536
; LENGTH: 171

```

Alignment Scores:		
Pred. No.:	1.99e-50	171
Score:	66.00	66
Percent Similarity:	100.00%	0
Best Local Similarity:	100.00%	0
Query Match:	5.89%	0
Conservative:	0	
Mismatches:	0	
Indels:	0	

DB: 9 Gaps: 0

US-09-270-437D-6 (1-3412) x US-09-764-864-1536 (1-171)

QY 1609 ATCAGAGTCCTCTTCCACAGCTGGCGGCGTGCAATGGCAAAGGTGGCAAGCCGTGAAC 1668
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 86 IleArgValProSerSerThrAlaGlyArgValIleGlyLysGlyGlyLysThrValAsn 105
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1669 GAATCTCAGAACAATTAAACCAGTGAGAACGATCATCTGTCGCTCGTAGCACCAAAAGCCGCAGATGAA 1728
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 106 GluLeuGlnAsnLeuThrSerAlaGluValIleValProArgAspGlnThrProAspGlu 125
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1729 AATCAGAGACTGATCTCTCAGAAATTATCGGACATTTCTTTGCTAGCCAGACTGCACAGCGC 1788
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 126 AsnGluGluValIleValArgIleIleGlyHisPhePheAlaSerGlnThrAlaGlnArg 145
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1789 AAGATCAGGGAATAATTGTA 1806
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 146 LysIleArgGluIleVal 151
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 8

US-10-262-445-40

; Sequence 40, Application US/10262445

; Publication No. US20040014058A1

GENERAL INFORMATION:

; APPLICANT: Alsbrook II, John

; APPLICANT: Burgess, Catherine

; APPLICANT: Catterton, Elina

; APPLICANT: Chant, John

; APPLICANT: Chaudhuri, Amitabha

; APPLICANT: Edinger, Shlomit

; APPLICANT: Gerlach, Valerie

; APPLICANT: Giot, Loic

; APPLICANT: Gorman, Linda

; APPLICANT: Guo, Xiaojia

; APPLICANT: Kekuda, Ramesh

; APPLICANT: Mezes, Peter

; APPLICANT: Millet, Isabelle

; APPLICANT: Ooi, Chean Eng

; APPLICANT: Patturajan, Meera

; APPLICANT: Rieger, Daniel

; APPLICANT: Spytek, Kimberly

; APPLICANT: Taupier Jr., Raymond J.

; APPLICANT: Zerhusen, Bryan

; APPLICANT: Zhong, Haihong

; APPLICANT: Zhong, Mei

TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND TITLE OF INVENTION: THE SAME

FILE REFERENCE: 21402-462D

CURRENT APPLICATION NUMBER: US/10/262,445

CURRENT FILING DATE: 2002-10-01

PRIOR APPLICATION NUMBER: 60/327,454

PRIOR FILING DATE: 2001-10-05

PRIOR APPLICATION NUMBER: 60/327,917

PRIOR FILING DATE: 2001-10-09

PRIOR APPLICATION NUMBER: 60/328,029

PRIOR FILING DATE: 2001-10-09

PRIOR APPLICATION NUMBER: 60/328,056

PRIOR FILING DATE: 2001-10-09

PRIOR APPLICATION NUMBER: 60/328,849

PRIOR FILING DATE: 2001-10-12

PRIOR APPLICATION NUMBER: 60/329,414

PRIOR FILING DATE: 2001-10-15

PRIOR APPLICATION NUMBER: 60/330,142

PRIOR FILING DATE: 2001-10-17

PRIOR APPLICATION NUMBER: 60/341,058

PRIOR FILING DATE: 2001-10-22

PRIOR APPLICATION NUMBER: 60/343,629

PRIOR FILING DATE: 2001-10-24

PRIOR APPLICATION NUMBER: 60/349,575

PRIOR FILING DATE: 2001-10-29

Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 133


```
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 40
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-262-445-40

Alignment Scores:
Pred. No.: 1.56e-30 Length: 555
Score: 44.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.93% Indels: 0
DB: 15 Gaps: 0

US-09-270-437D-6 (1-3412) x US-10-262-445-40 (1-555)

QY 325 ATTCGAAACATCCCTCTCCTACCTGCGAGTGTGGATGGACTTTGGCTCAATAT 384
Db 86 IleArgAsnIleProHisLeuGlnTrpGluValLeuAspGlyLeuLeuAlaGlnTyr 105
QY 385 GGCACAGTGGAGATGTGGACAAAGTCAACACACACAGACACAGAAACCGCGTTGTCAACGTC 444
Db 106 GlyThrValGluAsnValGluGlnValAsnThrAspThrGluThrAlaValValAsnVal 125
QY 445 ACATATGCAACA 456
Db 126 ThrTyrAlaThr 129

RESULT 9
US-10-313-986-499
; Sequence 499, Application US/10313986
; Publication No. US2003026209A1
; GENERAL INFORMATION:
; APPLICANT: FOY, Teresa M.
; APPLICANT: McNabb, Andria
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455G19
; CURRENT APPLICATION NUMBER: US/10/313,986
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 499
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-313-986-499

Alignment Scores:
Pred. No.: 1.71e-08 Length: 20
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.79% Indels: 0
DB: 15 Gaps: 0

US-09-270-437D-6 (1-3412) x US-10-313-986-499 (1-20)

QY 1453 AGATTGCCCTCGGAGAGCCAGAGCGTCAGGAAAGATGTCATCATCACCAGGCGCA 1512
Db 1 LysIleAlaProAlaGluGlyProAspValSerGluArgMetValIlelleThrGlyPro 20

RESULT 10
US-09-764-864-1534
; Sequence 1534, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PTZ23
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1534
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (24)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1534

Alignment Scores:
Pred. No.: 1.48e-08 Length: 47
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.79% Indels: 0
DB: 9 Gaps: 0

US-09-270-437D-6 (1-3412) x US-09-764-864-1534 (1-47)

QY 643 ATTGATTCCCGTCCGATCCTGTCGCCACCCAGTTTGTGGTCCCATCATCGGAAG 702
Db 1 IleAspPheProLeuArgIleLeuValProThrGlnPheValGlyAlaIlelleGlyLys 20

RESULT 11
US-09-873-637-20
; Sequence 20, Application US/09873637
; Patent No. US20020061543A1
; GENERAL INFORMATION:
; APPLICANT: Ross, Jeffrey
; TITLE OF INVENTION: THE C-MYC CODING REGION DETERMINANT-BINDING PROTEIN
; TITLE OF INVENTION: (CRD-BP) AND ITS NUCLEIC ACID SEQUENCE
; FILE REFERENCE: 960296.95131
; CURRENT APPLICATION NUMBER: US/09/873,637
; CURRENT FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-873-637-20

Alignment Scores:
Pred. No.: 1.2e-07 Length: 48
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.70% Indels: 0
DB: 9 Gaps: 0

US-09-270-437D-6 (1-3412) x US-09-873-637-20 (1-48)

QY 1630 GCTGGCGCGGTGATTGGCAAGGTGCGAGACCGTGAACGAACTGCAGAACTTAACC 1686
Db 8 AlaGlyArgValIleGlyLysGlyThrValAsnGluLeuGlnAsnLeuThr 26

RESULT 12
US-10-117-982-476
; Sequence 476, Application US/10117982
; Publication No. US20030138438A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
```

```
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Mericle, Barbara
; APPLICANT: Spies, Gregory A.
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C18
; CURRENT APPLICATION NUMBER: US/10/117,982
; CURRENT FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 484
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 476
; LENGTH: 81
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-117-982-476

Alignment Scores:
Pred. No.: 118-07 Length: 81
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.70% Indels: 0
DB: 14 Gaps: 0

US-09-270-437D-6 (1-3412) x US-10-117-982-476 (1-81)

Qy 1489 AGGATGTCATCATCACCGGCCCAAGCCAGTTCAGGCCCGGACGGATC 1545
Db 24 ArgMetValIleIleThrGlyProGluAlaGlnPheLysAlaGlnGlyArgIle 42

RESULT 13
US-10-313-986-476
; Sequence 476, Application US/10313986
; Publication No. US20030236209A1
; GENERAL INFORMATION:
; APPLICANT: Poy, Teresa M.
; APPLICANT: McNabb, Andria
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C19
; CURRENT APPLICATION NUMBER: US/10/313,986
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 476
; LENGTH: 81
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-313-986-476

Alignment Scores:
Pred. No.: 118-07 Length: 81
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.70% Indels: 0
DB: 15 Gaps: 0

US-09-270-437D-6 (1-3412) x US-10-313-986-476 (1-81)

Qy 1489 AGGATGTCATCATCACCGGCCCAAGCCAGTTCAGGCCCGGACGGATC 1545
Db 24 ArgMetValIleIleThrGlyProGluAlaGlnPheLysAlaGlnGlyArgIle 42

RESULT 14
US-09-764-864-1532
; Sequence 1532, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ23
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1532
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-864-1532

Alignment Scores:
Pred. No.: 9-07e-08 Length: 250
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.70% Indels: 0
DB: 9 Gaps: 0

US-09-270-437D-6 (1-3412) x US-09-764-864-1532 (1-250)

Qy 1630 GCTGGCCGGTGATTCGCAAGGTGGCAAGCCCTGCAAGCTGCAGACTTAACC 1686
Db 172 AlaGlyArgValIleIleGlyLysGlyGlyLysThrValAsnGluLeuGlnAsnLeuThr 190

RESULT 15
US-09-764-864-1114
; Sequence 1114, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ23
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1114
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-864-1114

Alignment Scores:
Pred. No.: 9e-08 Length: 261
Score: 19.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.70% Indels: 0
DB: 9 Gaps: 0

US-09-270-437D-6 (1-3412) x US-09-764-864-1114 (1-261)

Qy 1489 AGGATGTCATCATCACCGGCCCAAGCCAGTTCAGGCCCGGACGGATC 1545
Db 136 ArgMetValIleIleThrGlyProGluAlaGlnPheLysAlaGlnGlyArgIle 154

Search completed: July 16, 2004, 11:56:55
Job time : 174.5 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 21, 2004, 06:04:06 ; Search time 12960 Seconds
(without alignments)
11410.990 Million cell updates/sec

Title: US-09-270-437D-6
Perfect score: 3412
Sequence: 1 ggcagcgaggagcgagga.....aaccttgaaaatgtttattt 3412

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.ste.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pin.*
- 35: em.htg.rod.*
- 36: em.htg.mam.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3412	100.0	3412	6	AR171865	AR171865 Sequence
2	3412	100.0	3412	6	AR343074	AR343074 Sequence
3	3412	100.0	3412	6	BD209925	BD209925 Isolated
4	3125	91.6	3633	9	BC021290	BC021290 Homo sapi
5	2143	62.8	3283	6	AR171867	AR171867 Sequence
6	2143	62.8	3283	6	AR343076	AR343076 Sequence
7	2143	62.8	3283	6	BD209927	BD209927 Isolated
8	1977	57.9	2010	9	AF117107	AF117107 Homo sapi
9	1810	53.0	3667	9	AF057352	AF057352 Homo sapi
10	1478	43.3	2640	9	BSM806243	BSM806243 Homo sapi
c 11	1478	43.3	98945	9	AC016961	AC016961 Homo sapi
c 12	1296	38.0	208273	2	AC134942	AC134942 Homo sapi
c 13	586	17.2	50320	2	AC126373	AC126373 Homo sapi
14	567	16.6	710	6	BD221443	BD221443 Human gen
15	565	16.6	50320	2	AC126373	AC126373 Homo sapi
c 16	428	12.5	222876	2	AC068379	AC068379 Homo sapi
c 17	391	11.5	27352	2	AC126372	AC126372 Homo sapi
18	390	11.4	222876	2	AC068379	AC068379 Homo sapi
19	347	10.2	364	6	AX887402	AX887402 Sequence
20	347	10.2	364	6	BD027012	BD027012 Sequence
21	285	8.4	300	6	BD220404	BD220404 Human gen
22	264	7.7	264	6	AX209284	AX209284 Sequence
23	264	7.7	364	6	AX198757	AX198757 Sequence
24	263	7.7	263	6	AX209467	AX209467 Sequence
c 25	263	7.7	263	6	AX209697	AX209697 Sequence
26	263	7.7	378	6	AX198941	AX198941 Sequence
c 27	249	7.3	146376	9	AC009247	AC009247 Homo sapi
28	249	7.3	163157	9	AC108670	AC108670 Homo sapi
29	222	6.5	222	6	AX261025	AX261025 Sequence
30	222	6.5	222	6	AX261634	AX261634 Sequence
31	167	4.9	218	6	AX261764	AX261764 Sequence
32	156	4.6	306	6	BD059612	BD059612 Secreted
c 33	150	4.4	244	11	G43967	G43967 WIAF-3591-S
34	140	4.1	243	6	AX260963	AX260963 Sequence
35	135	4.0	142971	9	AC020629	AC020629 Homo sapi
36	127	3.7	3557	10	BC023758	BC023758 Mus muscu
37	127	3.7	4953	10	BC054552	BC054552 Mus muscu
38	100	2.9	91084	9	ALS96177	ALS96177 Human DNA
39	100	2.9	182695	2	AC015706	AC015706 Homo sapi
c 40	93	2.7	253	6	AX261078	AX261078 Sequence
c 41	92	2.7	67375	2	AC104767	AC104767 Homo sapi
c 42	92	2.7	112259	2	AC026579	AC026579 Homo sapi
c 43	92	2.7	113201	9	AP004230	AP004230 Homo sapi
c 44	92	2.7	187226	9	AC104980	AC104980 Homo sapi
c 45	83	2.4	380	6	AX335143	AX335143 Sequence

ALIGNMENTS

RESULT 1	AR171865	Sequence	6	3412 bp	DNA	linear	PAT 17-DEC-2001
LOCUS	AR171865	Sequence	6	from patent US 6297364.			
DEFINITION	AR171865	Sequence	6	from patent US 6297364.			
ACCESSION	AR171865	Sequence	6	from patent US 6297364.			
VERSION	AR171865.1	GI:17910815					
KEYWORDS	Unknown.						
SOURCE	Unknown.						
ORGANISM	Unknown.						
REFERENCE	1 (bases 1 to 3412)						
AUTHORS	Chen,Y.-T., Gure,A., Tsang,S., Stockert,E., Jager,E., Alexander,K.						
TITLE	Isolated nucleic acid molecule encoding cancer associated antigen, the antigen itself, and uses thereof						

JOURNAL Patent: US 6297364-A 6 02-OCT-2001; FEATURES Location/Qualifiers 1. .3412 source /organism="unknown" /mol_type="unassigned DNA"									
ORIGIN									
Query Match 100.0%; Score 3412; DB 6; Length 3412; Best Local Similarity 100.0%; Pred. No. 0; Matches 3412; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	GGCAGCGGAGAGGAGCGCGGTTACCGGCGCGGGGAGCGCGGCTCTCGGGG	60						
Db	1	GGCAGCGGAGAGGAGCGCGGTTACCGGCGCGGGGAGCGCGGCTCTCGGGG	60						
Qy	61	AAGAGACGGATGATGAACAAGCTTTACATCGGGAACCTGAGCCCGCGCTCACCGCGAC	120						
Db	61	AAGAGACGGATGATGAACAAGCTTTACATCGGGAACCTGAGCCCGCGCTCACCGCGAC	120						
Qy	121	GACCTCGGCGACTCTTTGGGGACAGGAAGCTGCCCCCTGGCGGACAGGTCTCTGAAG	180						
Db	121	GACCTCGGCGACTCTTTGGGGACAGGAAGCTGCCCCCTGGCGGACAGGTCTCTGAAG	180						
Qy	181	TCCGCTACGCTTCGTGACTACCCCGACAGAACTGGGCGCATCCGCGCCATCGAGACC	240						
Db	181	TCCGCTACGCTTCGTGACTACCCCGACAGAACTGGGCGCATCCGCGCCATCGAGACC	240						
Qy	241	CTCTCGGTAAGTGGAAATTCATGGGAAATCATGGAAGTTGATCTCAGTCTCTAAA	300						
Db	241	CTCTCGGTAAGTGGAAATTCATGGGAAATCATGGAAGTTGATCTCAGTCTCTAAA	300						
Qy	301	AAGCTAAGGAGAGGAAATTCAGATTGGAACATCCCTCTCAGCTCAGTGGAGGTG	360						
Db	301	AAGCTAAGGAGAGGAAATTCAGATTGGAACATCCCTCTCAGCTCAGTGGAGGTG	360						
Qy	361	TTGGATGACTTTTGCTCAATATGGGACAGTGGAGAAATGTTGATCTCAGTCTCTAAA	420						
Db	361	TTGGATGACTTTTGCTCAATATGGGACAGTGGAGAAATGTTGATCTCAGTCTCTAAA	420						
Qy	421	ACAGAAACCGCGTTGTCAACGTCAATATGGGACAGTGGAGAAATGTTGATCTCAGTCTCTAAA	480						
Db	421	ACAGAAACCGCGTTGTCAACGTCAATATGGGACAGTGGAGAAATGTTGATCTCAGTCTCTAAA	480						
Qy	481	GAGAGCTAAGCGGCGATCAGTTGAGAACTACTCTTCAAGATTCTCTACATCCCGAT	540						
Db	481	GAGAGCTAAGCGGCGATCAGTTGAGAACTACTCTTCAAGATTCTCTACATCCCGAT	540						
Qy	541	GAAGAGTGAAGTCCCTTTCGCGCTTCAAGCTCAATATGGGACAGTGGAGAAATGTTGATCTCAGTCTCTAAA	600						
Db	541	GAAGAGTGAAGTCCCTTTCGCGCTTCAAGCTCAATATGGGACAGTGGAGAAATGTTGATCTCAGTCTCTAAA	600						
Qy	601	GAGCAAGGCGAGCCCTTGGGGCACTTCTCAGGCGAGACAGATTGATTTCCCGTGGG	660						
Db	601	GAGCAAGGCGAGCCCTTGGGGCACTTCTCAGGCGAGACAGATTGATTTCCCGTGGG	660						
Qy	661	ATCCTGCTCCGACCCAGTTTGTGTCATCATCGGAAGAGGAGGCTTGACCATAAAG	720						
Db	661	ATCCTGCTCCGACCCAGTTTGTGTCATCATCGGAAGAGGAGGCTTGACCATAAAG	720						
Qy	721	AACATCACTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAAGAGAACTCTGGAGCT	780						
Db	721	AACATCACTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAAGAGAACTCTGGAGCT	780						
Qy	781	GCAGAGAGGCTGTCAACATCCATGTCACCCCGAGAGGGGACTTCTGAAGCATCCCGATG	840						
Db	781	GCAGAGAGGCTGTCAACATCCATGTCACCCCGAGAGGGGACTTCTGAAGCATCCCGATG	840						
Qy	841	ATTCTTGAATCATGCAAGAGGAGGAGGATGAGACCAAACTAGCCGAAGAGATTCCTCTG	900						
Db	841	ATTCTTGAATCATGCAAGAGGAGGAGGATGAGACCAAACTAGCCGAAGAGATTCCTCTG	900						
Qy	901	AAAATCTTGGCACAATGGCTTGGTTGGAAGCTGATTTGGAAGAAAGAGGCGAGAAATTG	960						

Db	901	AAAATCTTGGCACAATGGCTTGGTTGGAAGCTGATTTGGAAGAAAGAGGCGAGAAATTG	960						
Qy	961	AAGAAATTAAGATGAACAGGAGCAACAGATTAACATCTCATCTTTTGCAGGATTTGAGC	1020						
Db	961	AAGAAATTAAGATGAACAGGAGCAACAGATTAACATCTCATCTTTTGCAGGATTTGAGC	1020						
Qy	1021	ATATACAACCCGGAAAGAACCATCATCTGTGAAGGGCACTGTGAGGCTGTGCCAGTGTCT	1080						
Db	1021	ATATACAACCCGGAAAGAACCATCATCTGTGAAGGGCACTGTGAGGCTGTGCCAGTGTCT	1080						
Qy	1081	GAGATGAGATTAATGAAGAAAGCTGCTGAGGCGCTTTGAAAATGATGCTGGCTGTTAAC	1140						
Db	1081	GAGATGAGATTAATGAAGAAAGCTGCTGAGGCGCTTTGAAAATGATGCTGGCTGTTAAC	1140						
Qy	1141	CAACAAGCCATCTGATCCAGGCTTGAACCTCAGCGCACTTGGCACTTTTCAACAGGA	1200						
Db	1141	CAACAAGCCATCTGATCCAGGCTTGAACCTCAGCGCACTTGGCACTTTTCAACAGGA	1200						
Qy	1201	CTGTCCGTGCTATCTCCACAGCAGGCGCCCGGAGCTCCCGCTGCCCTTACCAC	1260						
Db	1201	CTGTCCGTGCTATCTCCACAGCAGGCGCCCGGAGCTCCCGCTGCCCTTACCAC	1260						
Qy	1261	CCCTTCACTACCCACTCCGGATCTTCCAGCCTGTACCCCTTACAGTTCGGCTTGGCCG	1320						
Db	1261	CCCTTCACTACCCACTCCGGATCTTCCAGCCTGTACCCCTTACAGTTCGGCTTGGCCG	1320						
Qy	1321	TTCCGCGATCATCTCTTATCCAGAGCAGGAGATTTGTAATCTCTTTCATCCCAACCCAG	1380						
Db	1321	TTCCGCGATCATCTCTTATCCAGAGCAGGAGATTTGTAATCTCTTTCATCCCAACCCAG	1380						
Qy	1381	GCTGTGGCGCCATCATCGGGAAGAGGGGCGACACATCAACAGCTGGCGAGATTCGCC	1440						
Db	1381	GCTGTGGCGCCATCATCGGGAAGAGGGGCGACACATCAACAGCTGGCGAGATTCGCC	1440						
Qy	1441	GGAGCTCTATCAAGATTGCCCCCTGCGAAGCGCCAGACGTCAGCGAAGAGATGTCATC	1500						
Db	1441	GGAGCTCTATCAAGATTGCCCCCTGCGAAGCGCCAGACGTCAGCGAAGAGATGTCATC	1500						
Qy	1501	ATCACCGGCGCCACCGGAAGCCAGTTCAAGGCGCCAGGACCGGATCTTTGGGAAACTGAAA	1560						
Db	1501	ATCACCGGCGCCACCGGAAGCCAGTTCAAGGCGCCAGGACCGGATCTTTGGGAAACTGAAA	1560						
Qy	1561	GAGGAAAACTTCTTTTAAACCCCAAGAAAGTGAAGCTGGAAGCGCATATCAGAGTGCC	1620						
Db	1561	GAGGAAAACTTCTTTTAAACCCCAAGAAAGTGAAGCTGGAAGCGCATATCAGAGTGCC	1620						
Qy	1621	TCCTTCCACAGCTGGCGGCTGATTTGGCAAGGTGCAAGCGTGAACCTGAGAGTGCAGAAC	1680						
Db	1621	TCCTTCCACAGCTGGCGGCTGATTTGGCAAGGTGCAAGCGTGAACCTGAGAGTGCAGAAC	1680						
Qy	1681	TTAACCCAGTGCAGAAAGTCACTGCTGCTGACCAAAAGCCAGATGAAAATGAGGAAGTG	1740						
Db	1681	TTAACCCAGTGCAGAAAGTCACTGCTGCTGACCAAAAGCCAGATGAAAATGAGGAAGTG	1740						
Qy	1741	ATCGTCAAGAAATATCGGGCACTTCTTTGCTAGCCAGACTGCAAGGCGAGATCAGGAA	1800						
Db	1741	ATCGTCAAGAAATATCGGGCACTTCTTTGCTAGCCAGACTGCAAGGCGAGATCAGGAA	1800						
Qy	1801	ATTGTACACAGCTGGAAGCAGGAGCAAGAAATACCTTCAGGGAGTCCCTTCAACAGCG	1860						
Db	1801	ATTGTACACAGCTGGAAGCAGGAGCAAGAAATACCTTCAGGGAGTCCCTTCAACAGCG	1860						
Qy	1861	AGCAAGTGAAGGCTTCCCAAGGCAAGGAGCAAGAAATACCTTCAGGGAGTCCCTTCAACAC	1920						
Db	1861	AGCAAGTGAAGGCTTCCCAAGGCAAGGAGCAAGAAATACCTTCAGGGAGTCCCTTCAACAC	1920						
Qy	1921	CTGACAGAAATGAGAACCGAGCGCAGATCGGAGCAAGCAAGACCATCTGAG	1980						
Db	1921	CTGACAGAAATGAGAACCGAGCGCAGATCGGAGCAAGCAAGACCATCTGAG	1980						
Qy	1981	GAAATGAGAGTCTGCGGAGCGCGCAGGAGCTCTGCGGAGGCGCTTGGAAACCCAGGGGC	2040						
Db	1981	GAAATGAGAGTCTGCGGAGCGCGCAGGAGCTCTGCGGAGGCGCTTGGAAACCCAGGGGC	2040						

```
QY 2041 CGAGGAGGGCGGGGAAGTCTAGCCAGGTTTGCAGAAACACGAGGCCCGGCTCCCGCC 2100
Db 2041 CGAGGAGGGCGGGGAAGTCTAGCCAGGTTTGCAGAAACACGAGGCCCGGCTCCCGCC 2100
QY 2101 CCCAGGGGCTTCTGAGGGCTTCAGGCATCCACTTCACCATCCACTCGGATCTCTCTGAA 2160
Db 2101 CCCAGGGGCTTCTGAGGGCTTCAGGCATCCACTTCACCATCCACTCGGATCTCTCTGAA 2160
QY 2161 CTCCACGACGCTATCCCTTTAGTTGAACATACTAGTGAACGCTGTTCAAGGCAAGC 2220
Db 2161 CTCCACGACGCTATCCCTTTAGTTGAACATACTAGTGAACGCTGTTCAAGGCAAGC 2220
QY 2221 AAAATGCACACCTTTTCTGTGGCAATCGTCTGTACATGTGTGACATATTAGAAA 2280
Db 2221 AAAATGCACACCTTTTCTGTGGCAATCGTCTGTACATGTGTGACATATTAGAAA 2280
QY 2281 GGAAGATGTTAAGATATGTGGCTGTGGTTTACACAGGGTCCCTGACGGGTATATAT 2340
Db 2281 GGAAGATGTTAAGATATGTGGCTGTGGTTTACACAGGGTCCCTGACGGGTATATAT 2340
QY 2341 TTAGAAATATATATCAATAACTCACTAACTCCAATTTTAAATCAATTATTTT 2400
Db 2341 TTAGAAATATATATCAATAACTCACTAACTCCAATTTTAAATCAATTATTTT 2400
QY 2401 TTTTCTTTTAAAGAGAGCAGGCTTTCTAGACTTTAAAGATAAAGTCTTTGGGAG 2460
Db 2401 TTTTCTTTTAAAGAGAGCAGGCTTTCTAGACTTTTAAAGATAAAGTCTTTGGGAG 2460
QY 2461 GTCTACGGGTAGAGAGGAGCTTTGAGGCCACCGCACAAAATTCACCCAGAGGAAT 2520
Db 2461 GTCTACGGGTAGAGAGGAGCTTTGAGGCCACCGCACAAAATTCACCCAGAGGAAT 2520
QY 2521 CTGCTCGGAAGGACACTCAAGCAGTTCTGATCACTCTGTGTATGTCAACAGAGGATA 2580
Db 2521 CTGCTCGGAAGGACACTCAAGCAGTTCTGATCACTCTGTGTATGTCAACAGAGGATA 2580
QY 2581 CGGTCTCTTGAAGAGGAACCTGTCTACTCTCTCATGCTGTCTAGCTATACACCAT 2640
Db 2581 CGGTCTCTTGAAGAGGAACCTGTCTACTCTCTCATGCTGTCTAGCTATACACCAT 2640
QY 2641 TCTCTTTGCTTCAACAGTTTAACTGTTTGTGATCTGATATATATTTCTGTCT 2700
Db 2641 TCTCTTTGCTTCAACAGTTTAACTGTTTGTGATCTGATATATATTTCTGTCT 2700
QY 2701 CTCTGTGTTATCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 2760
Db 2701 CTCTGTGTTATCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 2760
QY 2761 TTTTCTCATCCCTCCCTCATCCGATCTACGACCGCCCGCCCGCCCGCCCGCCCG 2820
Db 2761 TTTTCTCATCCCTCCCTCATCCGATCTACGACCGCCCGCCCGCCCGCCCGCCCG 2820
QY 2821 GTGCTCTGAGTATCATCACAAAAGGAAACAAAAGCGGAAACACACAAACAGCTCAA 2880
Db 2821 GTGCTCTGAGTATCATCACAAAAGGAAACAAAAGCGGAAACACACAAACAGCTCAA 2880
QY 2881 CTTACACTTGGTTACTCAAAAAGAAACAAAGTCAATGGTACTTGTCTAGCTTTTGAAG 2940
Db 2881 CTTACACTTGGTTACTCAAAAAGAAACAAAGTCAATGGTACTTGTCTAGCTTTTGAAG 2940
QY 2941 AGGAAACAGGAACCCACCAACCAACCAATCAACCAACCAAGAAAATTCACCAATG 3000
Db 2941 AGGAAACAGGAACCCACCAACCAACCAATCAACCAACCAAGAAAATTCACCAATG 3000
QY 3001 AAAGATGTTATTTGTCTTTTGTGATATAGCCATCAATATTCAGCAAAATG 3060
Db 3001 AAAGATGTTATTTGTCTTTTGTGATATAGCCATCAATATTCAGCAAAATG 3060
QY 3061 ATTCTCTTTTAAAAAATAATGTGGAGAAAGTAGAAATTTACCAAGTTGTGGCC 3120
Db 3061 ATTCTCTTTTAAAAAATAATGTGGAGAAAGTAGAAATTTACCAAGTTGTGGCC 3120
```

```
QY 3121 CAGGGCGTTAAATTCACAGATTTTTTTTAAACGAGAAAAACACACAGAAAGCTACCTCAG 3180
Db 3121 CAGGGCGTTAAATTCACAGATTTTTTTTAAACGAGAAAAACACACAGAAAGCTACCTCAG 3180
QY 3181 GTGTTTTTACCTCAGCACCTTCTCTGTGTCTTCCCTTAGAGATTTTGTAAAGCTCATAG 3240
Db 3181 GTGTTTTTACCTCAGCACCTTCTCTGTGTCTTCCCTTAGAGATTTTGTAAAGCTCATAG 3240
QY 3241 TTGGAGCATTTTTTTTTTAAATAAAATAGTTTGGAAAAAAATAAGATATCAACT 3300
Db 3241 TTGGAGCATTTTTTTTTTAAATAAAATAGTTTGGAAAAAAATAAGATATCAACT 3300
QY 3301 GCCAGCTTGGAGAGAGTGCAGTCCAGTGTGCAACAGCTGTTCTGAATGTCTCCGCT 3360
Db 3301 GCCAGCTTGGAGAGAGTGCAGTCCAGTGTGCAACAGCTGTTCTGAATGTCTCCGCT 3360
QY 3361 AGCCAAAGAACCNATATGCTTCTTTTGGACAAACCTTGGAAAAATGTTTATTT 3412
Db 3361 AGCCAAAGAACCNATATGCTTCTTTTGGACAAACCTTGGAAAAATGTTTATTT 3412

RESULT 2
AR343074
LOCUS AR343074 3412 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 6 from patent US 6576756.
ACCESSION AR343074
VERSION AR343074.1 GI:33738476
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3412)
AUTHORS Chen,Y.-T., Gure,A., Teang,S., Stockert,E., Jager,E., Alexander,K.
and Old,L.J.
TITLE Isolated nucleic acid molecule encoding cancer associated antigen,
the antigen itself, and uses thereof
Patent: US 6576756-A 6 10-JUN-2003;
JOURNAL Location/Qualifiers
FEATURES
source
1..3412
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 3412; DB 6; Length 3412;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3412; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCACGGAGGAGGAGGAGCGCGGTACCGGCGCGGGGAGCGCGGGCTCTCGGGG 60
Db 1 GGCACGGAGGAGGAGGAGCGCGGTACCGGCGCGGGGAGCGCGGGCTCTCGGGG 60
QY 61 AAGAGACGGATGATGAACAAGCTTTACATCGGGAACCTGAGCCCGCCGCTCACCGCCGAC 120
Db 61 AAGAGACGGATGATGAACAAGCTTTACATCGGGAACCTGAGCCCGCCGCTCACCGCCGAC 120
QY 121 GACCTCCGGCAGCTTTTGGGGAAGGAGTGCCTTGGCGGGGACAGGTCCTGTGAAG 180
Db 121 GACCTCCGGCAGCTTTTGGGGAAGGAGTGCCTTGGCGGGGACAGGTCCTGTGAAG 180
QY 181 TCCGCTACGCTTCTGTGACTACCCGACAGAACTCGGCCATCCGCGCCATCCAGACC 240
Db 181 TCCGCTACGCTTCTGTGACTACCCGACAGAACTCGGCCATCCGCGCCATCCAGACC 240
QY 241 CTCTCGGTTAAAGTGAATTCATGGGAAATTCATGGAATTTGATTCTCAGTCTCTAAA 300
Db 241 CTCTCGGTTAAAGTGAATTCATGGGAAATTCATGGAATTTGATTCTCAGTCTCTAAA 300
QY 301 AAGCTAAGAGCAGGAAATTCAGATTCGAAACATCCCTCTCCTCAGCTGAGGAGGTG 360
Db 301 AAGCTAAGAGCAGGAAATTCAGATTCGAAACATCCCTCTCCTCAGCTGAGGAGGTG 360
QY 361 TTGGATGACATTTTGGCTCAATATGGGACAGTGGAGATGTGGAACTTACACAGAC 420
Db 361 TTGGATGACATTTTGGCTCAATATGGGACAGTGGAGATGTGGAACTTACACAGAC 420
```

Db 361 TTGGATGGACCTTTTGGCTCAATATGGGACAGTGGAGAACTGGGAACAGTCAACACAGAC 420
Qy
421 ACAGAAACCCCGTGTGTCAACGTCACATATGCAACAAAGAGAAAGCAAAATAAGCCATG 480
Db
421 ACAGAAACCCCGTGTGTCAACGTCACATATGCAACAAAGAGAAAGCAAAATAAGCCATG 480
Qy
481 GAGAAGCTAAGCGGGCATAGTTGGAACTACTCTTCAAGATTTTCTACATCCCGGAT 540
Db
481 GAGAAGCTAAGCGGGCATAGTTGGAACTACTCTTCAAGATTTTCTACATCCCGGAT 540
Qy
541 GAAGAGGTAGCTCCCTTCGCCCCCTCAGCGAGCCAGCGTGGGGACCACTCTTCCCGG 600
Db
541 GAAGAGGTAGCTCCCTTCGCCCCCTCAGCGAGCCAGCGTGGGGACCACTCTTCCCGG 600
Qy
601 GAGCAAGCCACGGCCCTGGGGCACTTCTCAGGCCAGACAGATTGATTTCCCGCTCGG 660
Db
601 GAGCAAGCCACGGCCCTGGGGCACTTCTCAGGCCAGACAGATTGATTTCCCGCTCGG 660
Qy
661 ATCTGTGTCGCCACCCAGTTTGTGTGGTCCATCATCGGAAAGGAGGCTTGACCATAAAG 720
Db
661 ATCTGTGTCGCCACCCAGTTTGTGTGGTCCATCATCGGAAAGGAGGCTTGACCATAAAG 720
Qy
721 AACATCACTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAGAGAACTCTGGAGCT 780
Db
721 AACATCACTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAGAGAACTCTGGAGCT 780
Qy
781 GCAGAGAGCTGTCAACCTCCATGCCCCAGAGGGGACTTCTGAAGCATGCCGATG 840
Db
781 GCAGAGAGCTGTCAACCTCCATGCCCCAGAGGGGACTTCTGAAGCATGCCGATG 840
Qy
841 ATTCTTGAATCATGCAAAAGAGGCAAGATGAGACCAAACTAGCCGAAAGAGATTCTCTG 900
Db
841 ATTCTTGAATCATGCAAAAGAGGCAAGATGAGACCAAACTAGCCGAAAGAGATTCTCTG 900
Qy
901 AAAATCTTGCAACAAATGCTTGGTGGAGACTGATTGGAAGAGGAGGAGGAGGAGG 960
Db
901 AAAATCTTGCAACAAATGCTTGGTGGAGACTGATTGGAAGAGGAGGAGGAGGAGG 960
Qy
961 AAGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 1020
Db
961 AAGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 1020
Qy
1021 ATATACAAACCCGGAAGAAACATCATCTGTGAAGGCAAGTGTAGGCGCTGTGCAGTGT 1080
Db
1021 ATATACAAACCCGGAAGAAACATCATCTGTGAAGGCAAGTGTAGGCGCTGTGCAGTGT 1080
Qy
1081 GAGATAGATTATGAAGAGCTGCGTGAAGGCTTTGAAATGATATGCTGGCTGTTAAC 1140
Db
1081 GAGATAGATTATGAAGAGCTGCGTGAAGGCTTTGAAATGATATGCTGGCTGTTAAC 1140
Qy
1141 CAACAAGCCCAATCTGATCCAGAGTTGAACTCAGCGCACTTGGCATCTTTTCAACAGGA 1200
Db
1141 CAACAAGCCCAATCTGATCCAGAGTTGAACTCAGCGCACTTGGCATCTTTTCAACAGGA 1200
Qy
1201 CTGTCGCTGATCTCCACAGCAGGCCCCGCGGAGCTCCCCCGCTGCCCTACCA 1260
Db
1201 CTGTCGCTGATCTCCACAGCAGGCCCCGCGGAGCTCCCCCGCTGCCCTACCA 1260
Qy
1261 CCGTCACTACCCACTCCGATACTTCTCAGCGCTGTACCCCATCAGATTGGCCG 1320
Db
1261 CCGTCACTACCCACTCCGATACTTCTCAGCGCTGTACCCCATCAGATTGGCCG 1320
Qy
1321 TTCCCGCATCATCACTTTATCAGAGCAGGAGTTGTGAATCTCTTCATCCCAACCCAG 1380
Db
1321 TTCCCGCATCATCACTTTATCAGAGCAGGAGTTGTGAATCTCTTCATCCCAACCCAG 1380
Qy
1381 GCTGTGGGCGCATCATCTTATCAGAGCAGGAGTTGTGAATCTCTTCATCCCAACCCAG 1440
Db
1381 GCTGTGGGCGCATCATCTTATCAGAGCAGGAGTTGTGAATCTCTTCATCCCAACCCAG 1440
Qy
1441 GAGGCTCTATCAAGATTGCCCTGGGAAGGCCAGCGCTCAGCGAAAGGATGGTCATC 1500
Db
1441 GAGGCTCTATCAAGATTGCCCTGGGAAGGCCAGCGCTCAGCGAAAGGATGGTCATC 1500

Qy 1501 ATCACCGGGCCACCGGAAGCCCAAGTTCAGAGCCACGGGACGAGTCTTTTGGGAAACTGAAA 1560
Db 1501 ATCACCGGGCCACCGGAAGCCCAAGTTCAGAGCCACGGGACGAGTCTTTTGGGAAACTGAAA 1560
Qy 1561 GAGGAAACTTCTTTTAAACCCCAAGAGAGTGAAGCTGAAGCGCATATCAGAGTCC 1620
Db 1561 GAGGAAACTTCTTTTAAACCCCAAGAGAGTGAAGCTGAAGCGCATATCAGAGTCC 1620
Qy 1621 TCTTCCACAGCTGGCCGGGTGATTGGCAAAAGTGGCAAGACCGTGAACGAATGCAGAAC 1680
Db 1621 TCTTCCACAGCTGGCCGGGTGATTGGCAAAAGTGGCAAGACCGTGAACGAATGCAGAAC 1680
Qy 1681 TTAACCAAGTGCAGAGTCACTGTCCTCTGTGACCAACGCCAGATGAATAGGAAGTG 1740
Db 1681 TTAACCAAGTGCAGAGTCACTGTCCTCTGTGACCAACGCCAGATGAATAGGAAGTG 1740
Qy 1741 ATCGTCAGAAATTATCGGGCACTTCTTTGTAGCCAGACTGCACAGGCCAAGATCAGGAA 1800
Db 1741 ATCGTCAGAAATTATCGGGCACTTCTTTGTAGCCAGACTGCACAGGCCAAGATCAGGAA 1800
Qy 1801 ATTGTCAACAGGTGAAGCAGAGGACGAGAAATACCTCAGGGAGTTCCTCACAGCGC 1860
Db 1801 ATTGTCAACAGGTGAAGCAGAGGACGAGAAATACCTCAGGGAGTTCCTCACAGCGC 1860
Qy 1861 AGCAAGTGGCTCCACAGGACACGCAAAAACAAAGGATGAATGTAGCCCTTCCACAC 1920
Db 1861 AGCAAGTGGCTCCACAGGACACGCAAAAACAAAGGATGAATGTAGCCCTTCCACAC 1920
Qy 1921 CTGACAGATGAGACCAACCGCAGCAGCAGATCGGGAGCAAAACCAAGACCATCTGAG 1980
Db 1921 CTGACAGATGAGACCAACCGCAGCAGCAGATCGGGAGCAAAACCAAGACCATCTGAG 1980
Qy 1981 GAATGAGAGTCTGCGGAGCGCGCCAGGACTCTGCGAGGCGCTGAGAACCCACAGGGG 2040
Db 1981 GAATGAGAGTCTGCGGAGCGCGCCAGGACTCTGCGAGGCGCTGAGAACCCACAGGGG 2040
Qy 2041 CGAGGAGGCGCGGAGGAGTCAAGAGTTTGCAGAACCAACAGCGCCCTGAGAACCCAGG 2100
Db 2041 CGAGGAGGCGCGGAGGAGTCAAGAGTTTGCAGAACCAACAGCGCCCTGAGAACCCAGG 2100
Qy 2101 CCCAGGCTTCTGAGGCTTCAAGCCATCCACTTCAACCTCCACTCGGATCTCTCTGAA 2160
Db 2101 CCCAGGCTTCTGAGGCTTCAAGCCATCCACTTCAACCTCCACTCGGATCTCTCTGAA 2160
Qy 2161 CTCCACAGCGCTATCCCTTTTGTAGTTGAACCTAACATAGGTGAACCTGTTCAAGCC 2220
Db 2161 CTCCACAGCGCTATCCCTTTTGTAGTTGAACCTAACATAGGTGAACCTGTTCAAGCC 2220
Qy 2221 AAAATGCACACCTTTTCTGTGGCAATCGTCTCTGTACATGTGTACATATTAGAAA 2280
Db 2221 AAAATGCACACCTTTTCTGTGGCAATCGTCTCTGTGTACATGTGTACATATTAGAAA 2280
Qy 2281 GGAAGATGTTAAGATATGTGGCTGTGGTTTACAGGGTGCCTCAGCGGTAATATAT 2340
Db 2281 GGAAGATGTTAAGATATGTGGCTGTGGTTTACAGGGTGCCTCAGCGGTAATATAT 2340
Qy 2341 TTTAGAAATAATATCAATACTCACTAACTCCAAATTTTAAATCAATTAATTTT 2400
Db 2341 TTTAGAAATAATATCAATACTCACTAACTCCAAATTTTAAATCAATTAATTTT 2400
Qy 2401 TTTTCTTTTAAAGAGAGAGGCTTTTCTAGACTTTTAAAGATAAAGTCTTTGGGAG 2460
Db 2401 TTTTCTTTTAAAGAGAGAGAGGCTTTTCTAGACTTTTAAAGATAAAGTCTTTGGGAG 2460
Qy 2461 GTCTCAGGTGTAGAGAGGCTTTGAGGCCACCCGCAAAAATTCACAGAGGGAAT 2520
Db 2461 GTCTCAGGTGTAGAGAGGCTTTGAGGCCACCCGCAAAAATTCACAGAGGGAAT 2520
Qy 2521 CTGTCGGAAGGACACTCAGCGGCTTGTGATCTGATCTGATCTGATCTGATCTGATCTG 2580
Db 2521 CTGTCGGAAGGACACTCAGCGGCTTGTGATCTGATCTGATCTGATCTGATCTGATCTG 2580

Db 601 GAGCAAGCGCAGCCCTCGGGGCACTTCTCAGGCGCAGACAGATTGATTTCCCGCTGCGG 560
QY 661 ATCTGGTCCCAACCCAGTTTGTGTGCGCATCATCGGAAAGGAGGCTGTGACCAATAAG 720
Db 661 ATCTGGTCCCAACCCAGTTTGTGTGCGCATCATCGGAAAGGAGGCTGTGACCAATAAG 720
QY 721 AACATCACTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAAGAGAACTCTGAGCT 780
Db 721 AACATCACTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAAGAGAACTCTGAGCT 780
QY 781 GCAGAGAAGCTGTCAACCATCCATGCCACCCAGAGGGACTTCTCAAGCATGCCCATG 840
Db 781 GCAGAGAAGCTGTCAACCATCCATGCCACCCAGAGGGACTTCTGAGCATGCCCATG 840
QY 841 ATTCTTTGAATCATGAGAAAGAGGAGGATGAGACCAAACTAGCCGAGATTTGAGC 900
Db 841 ATTCTTTGAATCATGAGAAAGAGGAGGATGAGACCAAACTAGCCGAGATTTGAGC 900
QY 901 AAAATCTTGGCACACAAATGSCCTTGGTTGGAGACTGATTGGAAGAGGCAAAATTTG 960
Db 901 AAAATCTTGGCACACAAATGSCCTTGGTTGGAGACTGATTGGAAGAGGCAAAATTTG 960
QY 961 AAGAAAATTTGAAATGAAACAGGAGCAAGATTAACAACTCTCATCTTTGAGGATTTGAGC 1020
Db 961 AAGAAAATTTGAAATGAAACAGGAGCAAGATTAACAACTCTCATCTTTGAGGATTTGAGC 1020
QY 1021 ATATACAAACCCGAAAGAACCATCACTGTGAAGGACAGTTGAGGCTGTGCCAGTGCT 1080
Db 1021 ATATACAAACCCGAAAGAACCATCACTGTGAAGGACAGTTGAGGCTGTGCCAGTGCT 1080
QY 1081 GAGATAGAGATTATGAGAGCTGCGTGAGGCTTTGAAAATGATATGCTGGCTGTTAAC 1140
Db 1081 GAGATAGAGATTATGAGAGCTGCGTGAGGCTTTGAAAATGATATGCTGGCTGTTAAC 1140
QY 1141 CAACAAAGCCAACTGATCCAGGGTTGAACCTCAGCGCACTTGGCATCTTTCAACAGGA 1200
Db 1141 CAACAAAGCCAACTGATCCAGGGTTGAACCTCAGCGCACTTGGCATCTTTCAACAGGA 1200
QY 1201 CTGTCCGTGTATCTCAACAGAGGGCCCGGAGCTCCCGGTGCGCCCTACAC 1260
Db 1201 CTGTCCGTGTATCTCAACAGAGGGCCCGGAGCTCCCGGTGCGCCCTACAC 1260
QY 1261 CCCTTCACTACCACTCCGATATCTCCAGCTGTACCCCATCACCAAGTTTGGCCG 1320
Db 1261 CCCTTCACTACCACTCCGATATCTCCAGCTGTACCCCATCACCAAGTTTGGCCG 1320
QY 1321 TTCCCGCATCATCTCTTATCCAGAGCAGGAGATTGTGAATCTCTTCATCCCAACCCAG 1380
Db 1321 TTCCCGCATCATCTCTTATCCAGAGCAGGAGATTGTGAATCTCTTCATCCCAACCCAG 1380
QY 1381 GCTGTGGGCGCATCATCGGAGAGAGGGGGGCACATCAACAGCTGGCGAGATTGCC 1440
Db 1381 GCTGTGGGCGCATCATCGGAGAGAGGGGGGCACATCAACAGCTGGCGAGATTGCC 1440
QY 1441 GGAGCTCTATCAAGATTGCCCTGCGGAAGGCCAGAGCTCAGCGAAAGATGGTCACTC 1500
Db 1441 GGAGCTCTATCAAGATTGCCCTGCGGAAGGCCAGAGCTCAGCGAAAGATGGTCACTC 1500
QY 1501 ATCAACGGGCGCACCGGAAGCCAGTTCAAGGCCCAAGGACGATCTTTGGGAACTGAAA 1560
Db 1501 ATCAACGGGCGCACCGGAAGCCAGTTCAAGGCCCAAGGACGATCTTTGGGAACTGAAA 1560
QY 1561 GAGGAAACTCTTTTAACCCCAAGAGAGTGAAGCTGAAGCGCATATCAGAGTGCCC 1620
Db 1561 GAGGAAACTCTTTTAACCCCAAGAGAGTGAAGCTGAAGCGCATATCAGAGTGCCC 1620
QY 1621 TCTTCCAAGCTGSCGGGTGATTGGCAAAAGGTGGCAAGCCGTTGAACGAACTGCAGAAC 1680
Db 1621 TCTTCCAAGCTGSCGGGTGATTGGCAAAAGGTGGCAAGCCGTTGAACGAACTGCAGAAC 1680
QY 1681 TTAACAGTGCAGAGTCACTGTCCTGTGACCAACCGCCAGATGAAATGAGGAAGTG 1740

Db 1681 TTAACAGTGCAGAGTCACTGTCCTGTGACCAAAAGCCAGATGAAAATGAGGAAGTG 1740
QY 1741 ATCGTCAGAAATATCGGGCACTTCTTGTAGCCAGACTGACAGGCGCAAGATCAGGGAA 1800
Db 1741 ATCGTCAGAAATATCGGGCACTTCTTGTAGCCAGACTGACAGGCGCAAGATCAGGGAA 1800
QY 1801 ATTGTACAAACAGTGAAGCAGCAGGAGCAGAAATACCTTCAGGAGTCCGCTTCACAGCGC 1860
Db 1801 ATTGTACAAACAGTGAAGCAGCAGGAGCAGAAATACCTTCAGGAGTCCGCTTCACAGCGC 1860
QY 1861 AGCAAGTGAAGTCCCAACAGGCAACAGCAAAACAAACGATGAATGTAGCCCTTCCAACAC 1920
Db 1861 AGCAAGTGAAGTCCCAACAGGCAACAGCAAAACAAACGATGAATGTAGCCCTTCCAACAC 1920
QY 1921 CTGACAGATGAGACCAAAACGAGCAGCCAGATCGGAGCAAAACCAAGAGCAATCTGAG 1980
Db 1921 CTGACAGATGAGACCAAAACGAGCAGCCAGATCGGAGCAAAACCAAGAGCAATCTGAG 1980
QY 1981 GAATGAGAAAGTCTGCGGAGGCGCCAGGGACTCTGCGGAGGCGCTGAGAAACCCAGGGGC 2040
Db 1981 GAATGAGAAAGTCTGCGGAGGCGCCAGGGACTCTGCGGAGGCGCTGAGAAACCCAGGGGC 2040
QY 2041 CGAGGAGGCGCGGGAAGGTCAAGCCAGGTTTGCAGAACCAACCGAGCCCGCTCCCGCC 2100
Db 2041 CGAGGAGGCGCGGGAAGGTCAAGCCAGGTTTGCAGAACCAACCGAGCCCGCTCCCGCC 2100
QY 2101 CCCAGGGCTTCTGAGGCTTCAAGCCATCACTTCCACTCCACTCGGATCTCTCTGAA 2160
Db 2101 CCCAGGGCTTCTGAGGCTTCAAGCCATCACTTCCACTCCACTCGGATCTCTCTGAA 2160
QY 2161 CTCCCAACGAGCTATCCCTTTAGTTGAATCAATAGGTGAACCTGTTCAAAGCCAAAGC 2220
Db 2161 CTCCCAACGAGCTATCCCTTTAGTTGAATCAATAGGTGAACCTGTTCAAAGCCAAAGC 2220
QY 2221 AAAATGCAACCCCTTTTCTGCGGCAATCGTCTCTGTACATGTGTGATATTAATAAT 2280
Db 2221 AAAATGCAACCCCTTTTCTGCGGCAATCGTCTCTGTACATGTGTGATATTAATAAT 2280
QY 2281 GGAAGATGTTAAGATATGTGGCTGTGGGTTACACAGGTGCTGACGGGTAAATAT 2340
Db 2281 GGAAGATGTTAAGATATGTGGCTGTGGGTTACACAGGTGCTGACGGGTAAATAT 2340
QY 2341 TTTAGAAATAATATCAAAATCACTCAACTCACTCACTCACTCACTCACTCACTCACT 2400
Db 2341 TTTAGAAATAATATCAAAATCACTCAACTCACTCACTCACTCACTCACTCACTCACT 2400
QY 2401 TTTTCTTTTAAAGAGAGAGGCTTTTCTAGACTTTTAAAGAAATAAAGTCTTTGGGAG 2460
Db 2401 TTTTCTTTTAAAGAGAGAGGCTTTTCTAGACTTTTAAAGAAATAAAGTCTTTGGGAG 2460
QY 2461 GTCTCACGGTGTAGAGAGAGCTTTTGAAGCCACCCGACAAAATTCACCCAGAGGAAAT 2520
Db 2461 GTCTCACGGTGTAGAGAGAGCTTTTGAAGCCACCCGACAAAATTCACCCAGAGGAAAT 2520
QY 2521 CTGCTCGGAAGGACACTCAAGGAGTCTGGATCACTGTGTATGTCAACAGAGGGATA 2580
Db 2521 CTGCTCGGAAGGACACTCAAGGAGTCTGGATCACTGTGTATGTCAACAGAGGGATA 2580
QY 2581 CGCTCTCTTGAAGAGAACTCTGTCACTCTCTCACTCTCTCTCTCTCTCTCTCTCTCT 2640
Db 2581 CGCTCTCTTGAAGAGAACTCTGTCACTCTCTCACTCTCTCTCTCTCTCTCTCTCTCT 2640
QY 2641 TCTCTTTGCTTCAAGGTTTAAACTGGTTTTGATATCTGATATATTTCTGTCT 2700
Db 2641 TCTCTTTGCTTCAAGGTTTAAACTGGTTTTGATATCTGATATATTTCTGTCT 2700
QY 2701 CTCTGTGTTATCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 2760
Db 2701 CTCTGTGTTATCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 2760
QY 2761 TTTTCTCTATCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 2820
Db 2761 TTTTCTCTATCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 2820

RESULT 4
 BC021290
 LOCUS
 DEFINITION
 Homo sapiens IGF-II mRNA-binding protein 2, mRNA (cdna clone
 HMC:29539 IMAGE:5090334), complete cds.
 BC021290
 ACCESSION
 BC021290.2 GI:33878041
 MGC.
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 3633)
 Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
 Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
 Altschul,S.F., Zeeberg,H., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
 Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,F.,
 Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
 Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
 Carninci,P., Prange,C., Raja,S.S., Loquellano,N.A., Peters,G.J.,
 Abramson,R.D., Mullah,S.O., Bosak,S.A., McEwan,P.J.,
 McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
 Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
 Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
 Fahey,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S.,
 Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
 Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
 Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
 Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smallos,D.E.,
 Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 22388257
 12477932
 2 (bases 1 to 3633)
 Strausberg, R.
 Direct Submission
 Submitted (14-JAN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 On Aug 19, 2003 this sequence version replaced gi:18204200.
 Contact: MGC help desk
 Email: gcaps-E@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
 DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc.mgc@nih.gov
 Akhtar, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantirip, S., Thomas, P.J., Touchman, J.W., Tsurgou, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>
 Series: IRAL Plate: 39 Row: 0 Column: 21
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 27552765.
 Location/Qualifiers
 1..3633
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:29539 IMAGE:5090334"
 /tissue_type="Pancreas, epithelioid carcinoma"
 /clone_lib="NIH-MGC 42"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 1..3633
 /gene="IMP-2"
 /note="synonym: VICKZ2"
 /db_xref="LocusID:10644"
 68..1864
 /codon_start=1
 /product="IGF-II mRNA-binding protein 2"
 /protein_id="AAH21290.1"
 /db_xref="GI:18204201"
 /db_xref="LocusID:10644"
 /translation="MNKLYIGNLSPAVTADLRQLFGDKRLPLAGQVILKSGVAFVDY PQDNVRAIETLSGKVELHGMEDVYSKSLRSRKTQIRNIPLHQLWEVLGDLA QYGTVEVNEQVNTDTAVVNTVYATREAKTAMEKLSHQFENYSFKSLYIPDEVS SPSPPOAQRGSHSREQAGPGSTQARQIDFPIRLVPTQFVGAIIKGEGLTKNI TKQTQSRVDIHRKENSAAEKPVTHATPEGTSEACRMILEIMQKAEDETKLAEIPL KILAHNGLVRILGKEGNRLKIEHTGKTIITSLQDLSIYNPERTITVGTQVAPCA SAEIIMKLRKAFENDMLAVNQANLIPGLNSALGIFSTGLSVLSPAGRGAPPA APYHPTTHSGYFSSLYPHQGFPPHHHSYEQEIVNLFIPTQAVGAILGKGAKHIQ QLARFAGASIKTAPAEIGPVDSRSMVITGPPAQKFAQGRIFGKLKEENFFNKPFEVK LEAHIRVPSGTAGRVIGKGGTVKNEQLNTSAEIVPRDQTPDNEEVIVRIIGHFFA SQTQARKIETIVQVQKQEQKTPQGVASQSRK"
 74..280
 /note="RRM; Region: RNA recognition motif"
 /db_xref="CDD:smart00360"
 650..853
 misc_feature
 misc_feature

/note="KH; Region: K homology RNA-binding domain"		901		CTTGGCACACAAATGGCTTGGTTGGAAGACTGATTGGAAGAGGACGCAAAATTTGAAGAA	960
/db_xref="CDD:smart00322"		QY		AAATTGAACATGAAACAGGACCAAGATAAACATCTCATCTTTGACGATTTGAGCATATA	1025
1340. 1555		Db			1020
/note="KH; Region: K homology RNA-binding domain"		QY		AAATTGAACATGAAACAGGACCAAGATAAACATCTCATCTTTGACGATTTGAGCATATA	1085
/db_xref="CDD:smart00322"		Db			1080
ORIGIN		QY		CAACCCGGAAGAACCATCACTGTGAAGGGCACAGTTGAGGCTGTGCGAGTGTGAGAT	1145
Query Match 91.6%; Score 3125; DB 9; Length 3633;		Db			1140
Best Local Similarity 99.9%; Pred. No. 0;		QY		AGAGATTATGAAGAGCTGCGTGAGGCTTTGAAAATGATATGCTGGCTGTTAAACCAACA	1205
Matches 3365; Conservative 0; Mismatches 0; Indels 2; Gaps 2;		Db			1200
QY		QY		AGCCCAATCTGATCCAGGGTTGAACTCAGCGCATCTGGCATCTTTTCAACAGGACTGTC	1265
Db		Db			1260
QY		QY		AGCCCAATCTGATCCAGGGTTGAACTCAGCGCATCTGGCATCTTTTCAACAGGACTGTC	1325
Db		Db			1320
QY		QY		CGTGTCTATCTCCACAGCAGGGCCCGGGAGCTCCCGCTGCCCTTACACCCCTT	1385
Db		Db			1380
QY		QY		CGTGTCTATCTCCACAGCAGGGCCCGGGAGCTCCCGCTGCCCTTACACCCCTT	1445
Db		Db			1440
QY		QY		CGTGTCTATCTCCACAGCAGGGCCCGGGAGCTCCCGCTGCCCTTACACCCCTT	1505
Db		Db			1500
QY		QY		CGTGTCTATCTCCACAGCAGGGCCCGGGAGCTCCCGCTGCCCTTACACCCCTT	1565
Db		Db			1560
QY		QY		CGTGTCTATCTCCACAGCAGGGCCCGGGAGCTCCCGCTGCCCTTACACCCCTT	1625
Db		Db			1620
QY		QY		CGTGTCTATCTCCACAGCAGGGCCCGGGAGCTCCCGCTGCCCTTACACCCCTT	1685
Db		Db			1680
QY		QY		CGTGTCTATCTCCACAGCAGGGCCCGGGAGCTCCCGCTGCCCTTACACCCCTT	1745
Db		Db			1740
QY		QY		CGTGTCTATCTCCACAGCAGGGCCCGGGAGCTCCCGCTGCCCTTACACCCCTT	1805
Db		Db			1800
QY		QY		CGTGTCTATCTCCACAGCAGGGCCCGGGAGCTCCCGCTGCCCTTACACCCCTT	1865
Db		Db			1860
QY		QY		CGTGTCTATCTCCACAGCAGGGCCCGGGAGCTCCCGCTGCCCTTACACCCCTT	1925
Db		Db			1920
QY		QY		CGTGTCTATCTCCACAGCAGGGCCCGGGAGCTCCCGCTGCCCTTACACCCCTT	1985
Db		Db			1980
QY		QY		CGTGTCTATCTCCACAGCAGGGCCCGGGAGCTCCCGCTGCCCTTACACCCCTT	2045
Db		Db			2040

/note="KH; Region: K homology RNA-binding domain"		901		CTTGGCACACAAATGGCTTGGTTGGAAGACTGATTGGAAGAGGACGCAAAATTTGAAGAA	960
/db_xref="CDD:smart00322"		QY		AAATTGAACATGAAACAGGACCAAGATAAACATCTCATCTTTGACGATTTGAGCATATA	1025
1340. 1555		Db			1020
/note="KH; Region: K homology RNA-binding domain"		QY		AAATTGAACATGAAACAGGACCAAGATAAACATCTCATCTTTGACGATTTGAGCATATA	1085
/db_xref="CDD:smart00322"		Db			1080
ORIGIN		QY		CAACCCGGAAGAACCATCACTGTGAAGGGCACAGTTGAGGCTGTGCGAGTGTGAGAT	1145
Query Match 91.6%; Score 3125; DB 9; Length 3633;		Db			1140
Best Local Similarity 99.9%; Pred. No. 0;		QY		AGAGATTATGAAGAGCTGCGTGAGGCTTTGAAAATGATATGCTGGCTGTTAAACCAACA	1205
Matches 3365; Conservative 0; Mismatches 0; Indels 2; Gaps 2;		Db			1200
QY		QY		AGCCCAATCTGATCCAGGGTTGAACTCAGCGCATCTGGCATCTTTTCAACAGGACTGTC	1265
Db		Db			1260
QY		QY		AGCCCAATCTGATCCAGGGTTGAACTCAGCGCATCTGGCATCTTTTCAACAGGACTGTC	1325
Db		Db			1320
QY		QY		CGTGTCTATCTCCACAGCAGGGCCCGGGAGCTCCCGCTGCCCTTACACCCCTT	1385
Db		Db			1380
QY		QY		CGTGTCTATCTCCACAGCAGGGCCCGGGAGCTCCCGCTGCCCTTACACCCCTT	1445
Db		Db			1440
QY		QY		CGTGTCTATCTCCACAGCAGGGCCCGGGAGCTCCCGCTGCCCTTACACCCCTT	1505
Db		Db			1500
QY		QY		CGTGTCTATCTCCACAGCAGGGCCCGGGAGCTCCCGCTGCCCTTACACCCCTT	1565
Db		Db			1560
QY		QY		CGTGTCTATCTCCACAGCAGGGCCCGGGAGCTCCCGCTGCCCTTACACCCCTT	1625
Db		Db			1620
QY		QY		CGTGTCTATCTCCACAGCAGGGCCCGGGAGCTCCCGCTGCCCTTACACCCCTT	1685
Db		Db			1680
QY		QY		CGTGTCTATCTCCACAGCAGGGCCCGGGAGCTCCCGCTGCCCTTACACCCCTT	1745
Db		Db			1740
QY		QY		CGTGTCTATCTCCACAGCAGGGCCCGGGAGCTCCCGCTGCCCTTACACCCCTT	1805
Db		Db			1800
QY		QY		CGTGTCTATCTCCACAGCAGGGCCCGGGAGCTCCCGCTGCCCTTACACCCCTT	1865
Db		Db			1860
QY		QY		CGTGTCTATCTCCACAGCAGGGCCCGGGAGCTCCCGCTGCCCTTACACCCCTT	1925
Db		Db			1920
QY		QY		CGTGTCTATCTCCACAGCAGGGCCCGGGAGCTCCCGCTGCCCTTACACCCCTT	1985
Db		Db			1980
QY		QY		CGTGTCTATCTCCACAGCAGGGCCCGGGAGCTCCCGCTGCCCTTACACCCCTT	2045
Db		Db			2040

QY 2046 AGGGGGGGGAAGTTCAGCCAGGTTTTCAGAGAACCCAGAGCCCGGCTCCCGCCCCCA 2105
Db |||||
QY 2041 AGGGGGGGGAAGTTCAGCCAGGTTTTCAGAGAACCCAGAGCCCGGCTCCCGCCCCCA 2100
Db |||||
QY 2106 GGGCTTCTGCAGGCTTCAGCCATCCACTTCACCATCCACTCGGATCTCTCGAATCCC 2165
Db |||||
QY 2101 GGGCTTCTGCAGGCTTCAGCCATCCACTTCACCATCCACTCGGATCTCTCGAATCCC 2160
Db |||||
QY 2166 ACCAGCCTATCCCTTTTAGTTGAACCTAACTAGGTGAAGCTTTCAAGCAAGCAAAAT 2225
Db |||||
QY 2161 ACCAGCCTATCCCTTTTAGTTGAACCTAACTAGGTGAAGCTTTCAAGCAAGCAAAAT 2220
Db |||||
QY 2226 GCACACCCCTTTTCTGTGGCAATCGTCTCTGTACATGTGTACATATTAGAAAGGAA 2285
Db |||||
QY 2221 GCACACCCCTTTTCTGTGGCAATCGTCTCTGTACATGTGTACATATTAGAAAGGAA 2280
Db |||||
QY 2286 GATGTTAAGATATGTGGCTGTGGGTTACACAGGTCCTGTCAGCGGTAAATATTATAG 2345
Db |||||
QY 2281 GATGTTAAGATATGTGGCTGTGGGTTACACAGGTCCTGTCAGCGGTAAATATTATAG 2340
Db |||||
QY 2346 AAATAATATCAAAATACTCACTAACTCCAAATTTTAAATCAATTAATTTTATTTT 2405
Db |||||
QY 2341 AAATAATATCAAAATACTCACTAACTCCAAATTTTAAATCAATTAATTTTATTTT 2400
Db |||||
QY 2406 CTTTTTAAAGAGAACGAGCTTTCTAGACTTTTAAAGATTAAGTCTTTGGAGGTCTC 2465
Db |||||
QY 2401 CTTTTTAAAGAGAACGAGCTTTCTAGACTTTTAAAGATTAAGTCTTTGGAGGTCTC 2460
Db |||||
QY 2466 AGGTTGAGAGAGCTTTTGAAGCCACCCGACAAATTTACCCAGAGGGAATCTGCT 2525
Db |||||
QY 2461 AGGTTGAGAGAGCTTTTGAAGCCACCCGACAAATTTACCCAGAGGGAATCTGCT 2520
Db |||||
QY 2526 CGGAAGGACACTCAGGAGTTCGTGATCACTGATGTATGTCAACAGAGGATACCGTC 2585
Db |||||
QY 2521 CGGAAGGACACTCAGGAGTTCGTGATCACTGATGTATGTCAACAGAGGATACCGTC 2580
Db |||||
QY 2586 TCCTTTGAAGAGAAACTCTGTCACTCCCTCATCCCTGCTAGCTCATACACCAATTCCT 2645
Db |||||
QY 2581 TCCTTTGAAGAGAAACTCTGTCACTCCCTCATCCCTGCTAGCTCATACACCAATTCCT 2640
Db |||||
QY 2646 TTGCTTCAAGGTTTAACTGGTTTTGTGATCTGCTATATATTTCTGTCTCTCTC 2705
Db |||||
QY 2641 TTGCTTCAAGGTTTAACTGGTTTTGTGATCTGCTATATATTTCTGTCTCTCTC 2700
Db |||||
QY 2706 TGTATATCTCTCCCTCCCTCCCTCCCTCTTCTTCCATCTCTCTCTTTTGAATTTCC 2765
Db |||||
QY 2701 TGTATATCTCTCCCTCCCTCCCTCCCTCTTCTTCCATCTCTCTCTTTTGAATTTCC 2760
Db |||||
QY 2766 TCATCCCTCCATCTCAATCCCGTATCTACGACCCCGCCCGCCCGCCAGGCAAGGAGTCT 2825
Db |||||
QY 2761 TCATCCCTCCATCTCAATCCCGTATCTACGCA - CCCCCCCCCCCCCAGGCAAGGAGTCT 2819
Db |||||
QY 2826 CTGAGTATCATCATCAACAAAGGAACAAAGCGAAACACACAAACAGGCTCAACTTAC 2885
Db |||||
QY 2820 CTGAGTATCATCATCAACAAAGGAACAAAGCGAAACACACAAACAGGCTCAACTTAC 2879
Db |||||
QY 2886 ACTTGGTTACTCAAAAGAACAGTCAATGGTACTTGTCTAGGTTTGGAAAGAGAA 2945
Db |||||
QY 2880 ACTTGGTTACTCAAAAGAACAGTCAATGGTACTTGTCTAGGTTTGGAAAGAGAA 2939
Db |||||
QY 2946 AACAGGAACCCACCAACCAACCAATCAACCAACAAAGAAATTTCCCAATGAAGA 3005
Db |||||
QY 2940 AACAGGAACCCACCAACCAACCAATCAACCAACAAAGAAATTTCCCAATGAAGA 2999
Db |||||
QY 3006 ATGTATTTTGTCTTTTGTGATTTTGGTGTATAAGCCATCAATATTAGCAAAATGATTC 3065
Db |||||
QY 3000 ATGTATTTTGTCTTTTGTGATTTTGGTGTATAAGCCATCAATATTAGCAAAATGATTC 3059
Db |||||
QY 3066 TTTCTTT - AAAAAAATAATGTGAGGAAAGTAGAATTTACAGGTTTGTGGCCAGG 3124
Db |||||
QY 3060 TTTCTTTAAAAAATAATGTGAGGAAAGTAGAATTTTACCAAGGTTGTGGCCAGG 3119
Db |||||

QY 3125 GCCTTAAATTCACAGATTTTAAACGAGAAAAACACACAGAAAGCTACCTCAGGTGT 3184
Db |||||
QY 3120 GCCTTAAATTCACAGATTTTAAACGAGAAAAACACACAGAAAGCTACCTCAGGTGT 3179
Db |||||
QY 3185 TTTTACCTCAGCACCCTGCTTGTGTTCCCTTAGAGATTTTGTAAAGCTGATAGTTGG 3244
Db |||||
QY 3180 TTTTACCTCAGCACCCTGCTTGTGTTCCCTTAGAGATTTTGTAAAGCTGATAGTTGG 3239
Db |||||
QY 3245 AGCATTTTATTTTAAATAAATAAGTTGGAAAAAATAAGATATCAACTGCCA 3304
Db |||||
QY 3240 AGCATTTTATTTTAAATAAATAAGTTGGAAAAAATAAGATATCAACTGCCA 3299
Db |||||
QY 3305 GCTCGAGAAGGTGACAGTCCCAAGTGTGCAACAGCTGTTCTGAATGTTCTCCGCTAGCC 3364
Db |||||
QY 3300 GCTCGAGAAGGTGACAGTCCCAAGTGTGCAACAGCTGTTCTGAATGTTCTCCGCTAGCC 3359
Db |||||
QY 3365 AAGAACC 3371
Db |||||
QY 3360 AAGAACC 3366
Db |||||
RESULT 5
ARI171867
LOCUS ARI171867 3283 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 8 from patent US 6297364.
ACCESSION ARI171867
VERSION ARI171867.1 GI:17910817
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3283)
AUTHORS Chen,Y.-T., Gure,A., Tsang,S., Stockert,E., Jager,E., Alexander,K.
and Old,L.J.
TITLE Isolated nucleic acid molecule encoding cancer associated antigen,
the antigen itself, and uses thereof
JOURNAL Patent: US 6297364-A 8 02-OCT-2001;
FEATURES Location/Qualifiers
source 1..3283
/organism="unknown"
/mol_type="unassigned DNA"
Query Match 62.8%; Score 2143; DB 6; Length 3283;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1270 ACCCACTCCGATACCTTCTCCAGCTGTACCCCATCACCAAGTTTGGCCCGTTCCCGCAT 1329
Db |||||
QY 1141 ACCCACTCCGATACCTTCTCCAGCTGTACCCCATCACCAAGTTTGGCCCGTTCCCGCAT 1200
Db |||||
QY 1330 CATCACTCTTATCCAGAGCAGGAGATTGTGAATCTTTCATCCCAACCCAGGCTGTGGC 1389
Db |||||
QY 1201 CATCACTCTTATCCAGAGCAGGAGATTGTGAATCTTCTTCATCCCAACCCAGGCTGTGGC 1260
Db |||||
QY 1390 GCCATCATCGGAAAGAGGGGCAACATCAACAGCTGGCGAGATTGGCGGAGCTCT 1449
Db |||||
QY 1261 GCCATCATCGGAAAGAGGGGCAACATCAACAGCTGGCGAGATTGGCGGAGCTCT 1320
Db |||||
QY 1450 ATCAAGATTGCCCTCTCGGAAGGCCACAGCTCAGCGAAAGGATGGTCAATCATCACCGGG 1509
Db |||||
QY 1321 ATCAAGATTGCCCTCTCGGAAGGCCACAGCTCAGCGAAAGGATGGTCAATCATCACCGGG 1380
Db |||||
QY 1510 CCACCGAAGCCAGTTTCAGGCCCAGGACGGATCTTTGGGAAACTGAAAGAGGAAAC 1569
Db |||||
QY 1381 CCACCGAAGCCAGTTTCAGGCCCAGGACGGATCTTTGGGAAACTGAAAGAGGAAAC 1440
Db |||||
QY 1570 TTTCTTAAACCCCAAGAAAGTGAAGCTGGAAGCCATATCAGAGTGCCTCTTCCACA 1629
Db |||||
QY 1441 TTTCTTAAACCCCAAGAAAGTGAAGCTGGAAGCCATATCAGAGTGCCTCTTCCACA 1500
Db |||||
QY 1630 GCTGGCGGGTGAATGGCAAGGTGGCAAGCCGTGAACGAACTGCAGAACTTAAACAGT 1689
Db |||||

Db	1501	GCTGCCGGTGATTTGGCAAAAGGTGGCAAGACCGTGAAACGAACTGCAGAACTTAAACAGT	1560
Qy	1690	GCAGAAGTCACTCGTCCCTCGTGACCAAAAGCGCAGATGAAATAGAAAGTGAATCGTCAGA	1749
Db	1561	GCAAGTCACTCGTCCCTCGTGACCAAAAGCGCAGATGAAATAGAAAGTGAATCGTCAGA	1620
Qy	1750	ATTATCGGCACTCTTTGCTAGCCAGACTGCAAGCGCAAGATCAGGGAATTTGACAA	1809
Db	1621	ATTATCGGCACTCTTTGCTAGCCAGACTGCAAGCGCAAGATCAGGGAATTTGACAA	1680
Qy	1810	CAGTGAAGCAGCAGAGCAGAAATACCTCAGGAGTCGCTCAAGCGCAGCAAGTGA	1869
Db	1681	CAGTGAAGCAGCAGAGCAGAAATACCTCAGGAGTCGCTCAAGCGCAGCAAGTGA	1740
Qy	1870	GGCTCCCAAGCAGCAGCAAAACCAACGATGAATGTAGCCCTTCCAAACCTCTGACAGAA	1929
Db	1741	GGCTCCCAAGCAGCAGCAAAACCAACGATGAATGTAGCCCTTCCAAACCTCTGACAGAA	1800
Qy	1930	TGAGACCAAAACGAGCAGCAGATCGGAGCAAAACCAAGACCATCTCAGGAATGAGAA	1989
Db	1801	TGAGACCAAAACGAGCAGCAGATCGGAGCAAAACCAAGACCATCTCAGGAATGAGAA	1860
Qy	1990	GTCGCGAGGCGGCGAGGACTCTCCGAGGCGCTGAGAACCCCGAGGCGCGAGAGGG	2049
Db	1861	GTCGCGAGGCGGCGAGGACTCTCCGAGGCGCTGAGAACCCCGAGGCGCGAGAGGG	1920
Qy	2050	GCGGGGAAGGTGAGCGAGGTTTGCCAGAACCCAGGCGCGCTCCCGCGCCCGCAGGGC	2109
Db	1921	GCGGGGAAGGTGAGCGAGGTTTGCCAGAACCCAGGCGCGCTCCCGCGCCCGCAGGGC	1980
Qy	2110	TTCTGAGGCTTTCAGGCTCCTTCCATCCTCAGTCCGATCTCTCTGACCTCCCAAGCA	2169
Db	1981	TTCTGAGGCTTTCAGGCTCCTTCCATCCTCAGTCCGATCTCTCTGACCTCCCAAGCA	2040
Qy	2170	CGCTATCCCTTTAGTTGAATCACTAGTGAACGTTTCAAGGCAAGCAAAATGCAAC	2229
Db	2041	CGCTATCCCTTTAGTTGAATCACTAGTGAACGTTTCAAGGCAAGCAAAATGCAAC	2100
Qy	2230	ACCTTTTCTGTGGCAAAATCGTCTGTATCATGTGTGATCATATTAAGAGGAAGATG	2289
Db	2101	ACCTTTTCTGTGGCAAAATCGTCTGTATCATGTGTGATCATATTAAGAGGAAGATG	2160
Qy	2290	TTAAGATATGTGGCTGTGGGTTACACAGGTCCTGACGGTGAATATATTTAGAAAT	2349
Db	2161	TTAAGATATGTGGCTGTGGGTTACACAGGTCCTGACGGTGAATATATTTAGAAAT	2220
Qy	2350	AATATATCAATACTCAACTCACTCAATTTTAAATCAATTAATTTTCTTTTCTTTT	2409
Db	2221	AATATATCAATACTCAACTCACTCAATTTTAAATCAATTAATTTTCTTTTCTTTT	2280
Qy	2410	TTAAGAGAAAGCAGGCTTTTCTAGACTTTAAGAAATAAGTCTTTGGAGGTCACGG	2469
Db	2281	TTAAGAGAAAGCAGGCTTTTCTAGACTTTAAGAAATAAGTCTTTGGAGGTCACGG	2340
Qy	2470	TGTAGAGAGGCTTTGAGGCGCCAGCAGCAAAATTCACAGAGGAAATCTCGTCGGA	2529
Db	2341	TGTAGAGAGGCTTTGAGGCGCCAGCAGCAAAATTCACAGAGGAAATCTCGTCGGA	2400
Qy	2530	AGGACACTCAGCGAGTCTTGATCACCCTGTGTATGTCAACAGAGGATACCGTCTCCT	2589
Db	2401	AGGACACTCAGCGAGTCTTGATCACCCTGTGTATGTCAACAGAGGATACCGTCTCCT	2460
Qy	2590	TGAAGAGAAACTCTGTCTCCTCATGCTGTAGCTCATACACCAATTCCTTTGTC	2649
Db	2461	TGAAGAGAAACTCTGTCTCCTCATGCTGTAGCTCATACACCAATTCCTTTGTC	2520
Qy	2650	TTCAAGGCTTTTAAATCGTTTTTGGCATCTGCTATATATAATTTCTCTCTCTCTCTT	2709
Db	2521	TTCAAGGCTTTTAAATCGTTTTTGGCATCTGCTATATATAATTTCTCTCTCTCTCTT	2580
Qy	2710	TATCTCTCCCTCCCTCCCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2769
Db	2581	TATCTCTCCCTCCCTCCCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2640
Qy	2770	CCCTCCATCTCAATCCGTTATCTAGCACACCCCGCCCGCCCGCCCGCCCGCCCGCC	2829
Db	2641	CCCTCCATCTCAATCCGTTATCTAGCACACCCCGCCCGCCCGCCCGCCCGCCCGCC	2700
Qy	2830	GTATCACATCACAAAAAGGAACAAAGCGAAAAACACAAAAACCGCCTCACTTACATT	2889
Db	2701	GTATCACATCACAAAAAGGAACAAAGCGAAAAACACAAAAACCGCCTCACTTACATT	2760
Qy	2890	GGTTACTCAAAGAAACAAGAGTCAATGGTACTTGTCTAGCGTTTGGAAAGGAAACA	2949
Db	2761	GGTTACTCAAAGAAACAAGAGTCAATGGTACTTGTCTAGCGTTTGGAAAGGAAACA	2820
Qy	2950	GGAAACCCACCAACCAACCAATCAACCAACCAACCAACCAACCAACCAACCAACCA	3009
Db	2821	GGAAACCCACCAACCAACCAATCAACCAACCAACCAACCAACCAACCAACCAACCA	2880
Qy	3010	ATTTTGTCTTTTGTGATTTTGAAGCCATCAATTTTACCAAAATTTTCAAAATTTTCA	3069
Db	2881	ATTTTGTCTTTTGTGATTTTGAAGCCATCAATTTTACCAAAATTTTCAAAATTTTCA	2940
Qy	3070	TTTAAAAAAATTTTGAAGGAAAGTAAATTTTACCAAGTTTGTGGCCCGAGGCGTT	3129
Db	2941	TTTAAAAAAATTTTGAAGGAAAGTAAATTTTACCAAGTTTGTGGCCCGAGGCGTT	3000
Qy	3130	AAATTTCACAGATTTTAAACGAGAAAAACACACAGAAAGCTACCTCAGGTGTTTTTA	3189
Db	3001	AAATTTCACAGATTTTAAACGAGAAAAACACACAGAAAGCTACCTCAGGTGTTTTTA	3060
Qy	3190	CCTCAGCAGCTTGTCTTGTCTTCCCTTAGAGATTTTGAAGCTGATAGTTGGAGCAT	3249
Db	3061	CCTCAGCAGCTTGTCTTGTCTTCCCTTAGAGATTTTGAAGCTGATAGTTGGAGCAT	3120
Qy	3250	TTTTTTTATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	3309
Db	3121	TTTTTTTATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	3180
Qy	3310	GAGAAAGTCACTGCTCAAGTGTGCAACAGCTGTTCTGAAATGTTCTTCGCTAGCCAGAA	3369
Db	3181	GAGAAAGTCACTGCTCAAGTGTGCAACAGCTGTTCTGAAATGTTCTTCGCTAGCCAGAA	3240
Qy	3370	CCNATATGGCTTCTTTTGGACAAACCTTGAAATGTTTTTTT 3412	
Db	3241	CCNATATGGCTTCTTTTGGACAAACCTTGAAATGTTTTTTT 3283	
RESULT 6			
AR343076			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 2143; Conservative			
Qy			

3283 bp DNA linear PAT 17-AUG-2003
Sequence 8 from patent US 6576756.
AR343076
AR343076
AR343076.1 GI:33738478
Unknown.
Unclassified.
1 (bases 1 to 3283)
Chen, Y.-T., Gure, A., Tsang, S., Stockert, E., Jager, E., Alexander, K.
and Old, L.J.
isolated nucleic acid molecule encoding cancer associated antigen,
the antigen itself, and uses thereof
Patent: US 6576756-A 8 10-JUN-2003;
Location/Qualifiers
1..3283
/organism="unknown"
/mol_type="genomic DNA"

62.8%; Score 2143; DB 6; Length 3283;
100.0%; Pred. No. 0;
0; Mismatches 0; Indels 0; Gaps 0;
1270 ACCCACTCCGATACCTTCCAGCTGTATACCCCATCACCAGTTTGGCCGCTTCCGCAT 1329

|||||
1141 ACCCACTCGGATACCTCTCCAGCTGTACCCCATACCACTTTGGCCGCTCCGCGAT 1200
QY CATCACTTTATCCAGACGAGAGATTGTGAATCTCTTCATCCCAACCCAGGCTGTGGGC 1389
Db CATCACTTTATCCAGACGAGAGATTGTGAATCTCTTCATCCCAACCCAGGCTGTGGGC 1260
QY GCATCATCGGAAGAGAGAGGCGACACATCAACAGCTGGCGAGATTGCGCGGAGGCTCT 1449
Db GCCATCATCGGAAGAGAGAGGCGCACATCAACAGCTGGCGAGATTGCGCGGAGGCTCT 1320
QY ATCAAGATTGCCCTTGGCGAAGGCCAGACGTCAGCGAAAGGATGGTCAATCATCACCGGG 1509
Db ATCAAGATTGCCCTTGGCGAAGGCCAGACGTCAGCGAAAGGATGGTCAATCATCACCGGG 1380
QY COACCGGAAGCCAGTTCAAGGCCCCAGGACGGATCTTTGGAAACTGAAAGAGGAAAC 1569
Db CCACCGGAAGCCAGTTCAAGGCCCCAGGACGGATCTTTGGAAACTGAAAGAGGAAAC 1440
QY TTCCTTTAAACCCCAAGAGAGAGTGAAGCTGGAGGCGCATATCAGAGTGCCCTCTTCCACA 1629
Db TTCCTTTAAACCCCAAGAGAGAGTGAAGCTGGAGGCGCATATCAGAGTGCCCTCTTCCACA 1500
QY GTGCGCCGGGTGATTGGCAAGGTGGCAAGCCGTGAACGAACTGCAGAACTTAACCCAGT 1689
Db GTGCGCCGGGTGATTGGCAAGGTGGCAAGCCGTGAACGAACTGCAGAACTTAACCCAGT 1560
QY GGAGAAGTCATGTCGCTCGTGACCAAAACCGCAGATGAAATGAGGAAAGTATGTCAGA 1749
Db GCAGAAGTCATGTCGCTCGTGACCAAAACCGCAGATGAAATGAGGAAAGTATGTCAGA 1620
QY ATTATCGGCACTTCTTTGCTAGCCAGACTGCACAGCGCAAGATCAGGGAATTTGTACAA 1809
Db ATTATCGGCACTTCTTTGCTAGCCAGACTGCACAGCGCAAGATCAGGGAATTTGTACAA 1680
QY CAGGTGAAGCAGCAGGACGAGAAATACCTCAGGAGTGCCTCAGCGCAGCAAGTGA 1869
Db CAGGTGAAGCAGCAGGACGAGAAATACCTCAGGAGTGCCTCAGCGCAGCAAGTGA 1740
QY GGCTCCCAAGCAACAGCAAAACAAACCGATGAATGTAGCCCTTCCACACCTGCAGAA 1929
Db GGCTCCCAAGCAACAGCAAAACAAACCGATGAATGTAGCCCTTCCACACCTGCAGAA 1800
QY TGAGACCAAAACGACGCCAGCAGATCGGAGCAAAACCAAGCAATCTGAGGAATGAGAA 1989
Db TGAGACCAAAACGACGCCAGCAGATCGGAGCAAAACCAAGCAATCTGAGGAATGAGAA 1860
QY GTCTCGGAGGCGGCAGGCACTCTGCCGAGGCCCTGAGAACCCAGGCGCGAGGAGG 2049
Db GTCTCGGAGGCGGCAGGCACTCTGCCGAGGCCCTGAGAACCCAGGCGCGAGGAGG 1920
QY GCGGGGAAGGTGAGCCAGGTTTGCAGAACCAACGAGGCCCGCCCTCCGCCCCCAGGCG 2109
Db GCGGGGAAGGTGAGCCAGGTTTGCAGAACCAACGAGGCCCGCCCTCCGCCCCCAGGCG 1980
QY TTCCTCAGGCTTCAGCCATCCATTCACATCCATCGGATCTCTCCTGAACTCCACGA 2169
Db TTCCTCAGGCTTCAGCCATCCATTCACATCCATCGGATCTCTCCTGAACTCCACGA 2040
QY CGCTATCCCTTTAGTTGAACCTTAACATAGGTGAACGTTTCAAGCCAGCAAGCAAAATGCAC 2229
Db CGCTATCCCTTTAGTTGAACCTTAACATAGGTGAACGTTTCAAGCCAGCAAGCAAAATGCAC 2100
QY ACCCTTTTCTGTGGCAATCGTCTGTGTACATGTGTGTACATATTAGAAAGGGAAGATG 2289
Db ACCCTTTTCTGTGGCAATCGTCTGTGTACATGTGTGTACATATTAGAAAGGGAAGATG 2160
QY TTAAGATATGTGGCTGTGGGTTCACAGGTCCTGTGAGCGGTAAATATATTTTGAAT 2349
Db TTAAGATATGTGGCTGTGGGTTCACAGGTCCTGTGAGCGGTAAATATATTTTGAAT 2220
QY AATATATCAAAATACTCAACTAACTCCAAATTTTAAATCAATTAATTAATTTTCTTT 2409
|||||

2221 AATATATCAAAATACTCAACTAACTCCAAATTTTAAATCAATTAATTAATTTTCTTT 2280
QY TTAAGAGAGAAAGCAGGCTTTTCTAGACTTTAAAGAAATAAAGTCTTTGGGAGGTCTCACGG 2469
Db TTAAGAGAGAAAGCAGGCTTTTCTAGACTTTAAAGAAATAAAGTCTTTGGGAGGTCTCACGG 2340
QY TGTAGAGAGAGGCTTTGAGGCGCACCCGACAAAATTCACCCAGAGGGAATCTCGTCGA 2529
Db TGTAGAGAGAGGCTTTGAGGCGCACCCGACAAAATTCACCCAGAGGGAATCTCGTCGA 2400
QY AGGACACTCAGGCACTGTCGATCACTGTCATCACTGTCATACACCCATTTCTCTTTC 2589
Db AGGACACTCAGGCACTGTCGATCACTGTCATCACTGTCATACACCCATTTCTCTTTC 2460
QY TGAAGAGGAAACTCTGTCTCACTCCTCATGCTGTCTAGCTCATACACCCATTTCTCTTTC 2649
Db TGAAGAGGAAACTCTGTCTCACTCCTCATGCTGTCTAGCTCATACACCCATTTCTCTTTC 2520
QY TTCACAGGTTTAAACTGGTTTTTGGCATACCTGCTATATAAATCTCTCTCTCTCTCT 2709
Db TTCACAGGTTTAAACTGGTTTTTGGCATACCTGCTATATAAATCTCTCTCTCTCTCTCT 2580
QY TATCTCTCCCTCCCTCCCTCCCTCTCTCCATCTCCATCTCTTTGAAATTCCTCAT 2769
Db TATCTCTCCCTCCCTCCCTCCCTCTCTCCATCTCCATCTCTTTGAAATTCCTCAT 2640
QY CCTCATCTCAATCCCTGCTATCTAGCACCCCTCCCTCCCTCCCTCCCTGCTCTCTCT 2829
Db CCTCATCTCAATCCCTGCTATCTAGCACCCCTCCCTCCCTCCCTCCCTGCTCTCTCT 2700
QY GTATCATATCACAAAAGGAAACAAAAGGAAACACAAAACAGGCTCAACTTACAT 2889
Db GTATCATATCACAAAAGGAAACAAAAGGAAACACAAAACAGGCTCAACTTACAT 2760
QY GTTACTCAAAAGAACAAAGAGTCAATGTACTGTCTTAGCGTTTGGAGAGGAAACA 2949
Db GTTACTCAAAAGAACAAAGAGTCAATGTACTGTCTTAGCGTTTGGAGAGGAAACA 2820
QY GGAACCCCAACCAACCAATCAACCAAAACAAAGAAATAATTCACAATGAAAGATGT 3009
Db GGAACCCCAACCAACCAATCAACCAAAACAAAGAAATAATTCACAATGAAAGATGT 2880
QY ATTTTGTCTTTTGCATTTTGGTGTATAGGCATCAATATTCAGCAAAATGATTCCTTTC 3069
Db ATTTTGTCTTTTGCATTTTGGTGTATAGGCATCAATATTCAGCAAAATGATTCCTTTC 2940
QY TTTAAAAAAAATGTGGAGAAAGTAGAAATTTACCAAGTGTGTGCGCCAGGCGGT 3129
Db TTTAAAAAAAATGTGGAGAAAGTAGAAATTTACCAAGTGTGTGCGCCAGGCGGT 3000
QY TTTAAAAAAAATGTGGAGAAAGTAGAAATTTACCAAGTGTGTGCGCCAGGCGGT 3189
Db TTTAAAAAAAATGTGGAGAAAGTAGAAATTTACCAAGTGTGTGCGCCAGGCGGT 3060
QY CCTCAGCACTTGTCTTGTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCT 3249
Db CCTCAGCACTTGTCTTGTGTCTTGTGTCTTGTCTTGTCTTGTCTTGTCTTGTCTTGTCT 3120
QY TTTTATTTTAAATAAATAAGTGTGAAAAATAAGATATCACTGCCAGCCTG 3309
Db TTTTATTTTAAATAAATAAGTGTGAAAAATAAGATATCACTGCCAGCCTG 3180
QY GAGAAGGTGACAGTCCAAGTGTGCAACAGCTGTCTGAATTTGTCTTCCGCTAGCAAGAA 3369
Db GAGAAGGTGACAGTCCAAGTGTGCAACAGCTGTCTGAATTTGTCTTCCGCTAGCAAGAA 3240
QY CCNATATGGCCCTCTTTTGGCAAAACCTTGAAAAATGTTTATTT 3412
Db CCNATATGGCCCTCTTTTGGCAAAACCTTGAAAAATGTTTATTT 3283

LOCUS BD209927 3283 bp DNA linear PAT 17-JUL-2003
 DEFINITION Isolated nucleic acid molecules encoding cancer-associated
 antigens, these antigens and method of using the same.
 ACCESSION BD209927
 VERSION BD209927.1 GI:33019697
 KEYWORDS JP 2002512049-A/6.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 3283)
 AUTHORS Chen,Y.T., Gure,A., Teang,S., Stockert,E., Jager,E., Knuth,A. and
 Old,L.J.
 TITLE Isolated nucleic acid molecules encoding cancer-associated
 antigens, these antigens and method of using the same
 JOURNAL Patent: JP 2002512049-A 6 23-APR-2002;
 COMMENT LUDWIG INSTITUTE FOR CANCER RESEARCH
 OS Homo sapiens (human)
 PN JP 2002512049-A/6
 PD 23-APR-2002
 PF 16-MAR-1999 JP 2000545030
 PR 17-APR-1998 US 09/061709
 PI YAO TSENG CHEN,ALI GURE,SOLAM TSANG,ELISABETH STOCKERT,ELKE
 JAGER,
 PI ALEXANDER KNUTH,LLOYD J OLD
 PC C12N15/09,A61K35/12,A61K39/00,A61K39/39,A61P35/00,C07K16/32,
 C12N1/15,
 PC C12N1/19,C12N1/21,C12N5/10,C12P21/08,C12Q1/68,G01N33/53,G01N33/PC
 574,
 PC C12N15/00,C12N5/00
 CC n=not determined(3243)
 FH Key Location/Qualifiers
 FT source 1..3283
 FT Location/Qualifiers
 1..3283
 /organism="Homo sapiens (human)"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Query Match 62.8%; Score 2143; DB 6; Length 3283;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1270 ACCACATCCGGATACCTTCACGCTGTACCCCATCACCAGTTGGCCGTTCCCGCAT 1329
 DB 1141 ACCACATCCGGATACCTTCACGCTGTACCCCATCACCAGTTGGCCGTTCCCGCAT 1200
 QY 1330 CATCACTCTTATCCAGAGCAGGATGTGAATCTCTTATCCCAACCCAGGCTGTGGC 1389
 DB 1201 CATCACTCTTATCCAGAGCAGGATGTGAATCTCTTATCCCAACCCAGGCTGTGGC 1260
 QY 1390 GCCATCATCGGGAAGAGGGGACACATCAAAACAGCTGGCGAGATTTCGCGGAGCCTCT 1449
 DB 1261 GCCATCATCGGGAAGAGGGGACACATCAAAACAGCTGGCGAGATTTCGCGGAGCCTCT 1320
 QY 1450 ATCAAGATTGCCCTCGGGAAGGCCAGAGCTGAGGAAAGGATGGTATCATCAACCGG 1509
 DB 1321 ATCAAGATTGCCCTCGGGAAGGCCAGAGCTGAGGAAAGGATGGTATCATCAACCGG 1380
 QY 1510 CCACCGGAAGCCAGTTCAAGGGCCAGGAGCGGATCTTTGGGAAACTGAAAGAGGAAAC 1569
 DB 1381 CCACCGGAAGCCAGTTCAAGGGCCAGGAGCGGATCTTTGGGAAACTGAAAGAGGAAAC 1440
 QY 1570 TTCTTTAAACCCCAAGGAAGTGAAGCTGGAAGCGCATATCAAGTGCCTCTTCCACA 1629
 DB 1441 TTCTTTAAACCCCAAGGAAGTGAAGCTGGAAGCGCATATCAAGTGCCTCTTCCACA 1500
 QY 1630 GCTGGCCGGGTGATTGGCAAGGTGGCAAGCCGTGACCACTGACGAACCTTAACAGT 1689
 DB 1501 GCTGGCCGGGTGATTGGCAAGGTGGCAAGCCGTGACCACTGACGAACCTTAACAGT 1560

QY 1690 GCAGAAAGTATCGTGGCTGTGTGACCAAAACCCAGATGAAAATGAGGAAGTATGTCGAGA 1749
 DB 1561 GCAGAAAGTATCGTGGCTGTGTGACCAAAACCCAGATGAAAATGAGGAAGTATGTCGAGA 1620
 QY 1750 ATTATCGGGCACATCTTTTGTAGCCAGATGTGCACAGCGAAGATCAGGGAATTTGTACAA 1809
 DB 1621 ATTATCGGGCACATCTTTTGTAGCCAGATGTGCACAGCGCAAGATCAGGGAATTTGTACAA 1680
 QY 1810 CAGGTGAAGCAGCAGGAGCAGAAATACCTTCAGGGAGTGGCTTCACAGGCGCAGCAAGTGA 1869
 DB 1681 CAGGTGAAGCAGCAGGAGCAGAAATACCTTCAGGGAGTGGCTTCACAGGCGCAGCAAGTGA 1740
 QY 1870 GGTCTCCACAGGCACGACGAGCAAAACCAACGAGTGAATGTAGCCCTTCCAAACCTTCAGAGAA 1929
 DB 1741 GGTCTCCACAGGCACGAGCAAAACCAACGAGTGAATGTAGCCCTTCCAAACCTTCAGAGAA 1800
 QY 1930 TGAGACCAAAACGAGCAGCAGCAGATCGGAGCAAAACCAAGACCATCTGAGGAATGAGAA 1989
 DB 1801 TGAGACCAAAACGAGCAGCAGCAGATCGGAGCAAAACCAAGACCATCTGAGGAATGAGAA 1860
 QY 1990 GTCTGCGAGGCGGCGCAGGAGCTCTGCGAGGCGCTTGAGAAACCCAGGGGCCGAGGAGGG 2049
 DB 1861 GTCTGCGAGGCGGCGCAGGAGCTCTGCGAGGCGCTTGAGAAACCCAGGGGCCGAGGAGGG 1920
 QY 2050 GCGGGGAAGGTGACGAGGTTTGGCCAGAAACCCAGAGCCCGCTCCCGCCCCCAGGGC 2109
 DB 1921 GCGGGGAAGGTGACGAGGTTTGGCCAGAAACCCAGAGCCCGCTCCCGCCCCCAGGGC 1980
 QY 2110 TTCTGCAGGCTTCAGGCATCCACTTCACCATCCACTCGGATCTCTCTGAACCTCCACAGA 2169
 DB 1981 TTCTGCAGGCTTCAGGCATCCACTTCACCATCCACTCGGATCTCTCTGAACCTCCACAGA 2040
 QY 2170 CGCTATCCCTTTTAGTTGAACATAGTAGTAACGTGTGTACATATTAGAAAGGGAAGATG 2229
 DB 2041 CGCTATCCCTTTTAGTTGAACATAGTAGTAACGTGTGTACATATTAGAAAGGGAAGATG 2100
 QY 2230 ACCCTTTTCTGTGGCAAAATCGTCTGTGTACATATTAGAAAGGGAAGATG 2289
 DB 2101 ACCCTTTTCTGTGGCAAAATCGTCTGTGTACATATTAGAAAGGGAAGATG 2160
 QY 2290 TTAAGATATGTGCCTGTGGGTTTACACAGGGTGCCTGCAGCGGTAATATATTTAGAAAT 2349
 DB 2161 TTAAGATATGTGCCTGTGGGTTTACACAGGGTGCCTGCAGCGGTAATATATTTAGAAAT 2220
 QY 2350 AATATATCAATTAACATCAACTCAATTTTAACTTAATCAATTAATTTTCTTTT 2409
 DB 2221 AATATATCAATTAACATCAACTCAATTTTAACTTAATTAATTTTCTTTT 2280
 QY 2410 TTAAGAGAAAGCAGGCTTTCTAGACTTTTAAAGAAATAAGTCTTTGGGAGGTCTCACGG 2469
 DB 2281 TTAAGAGAAAGCAGGCTTTCTAGACTTTTAAAGAAATAAGTCTTTGGGAGGTCTCACGG 2340
 QY 2470 TGTAGAGAGAGCTTTGAGGCCACCCGCAAAAATTCACCCAGAGGAAATCTCGTGGGA 2529
 DB 2341 TGTAGAGAGAGCTTTGAGGCCACCCGCAAAAATTCACCCAGAGGAAATCTCGTGGGA 2400
 QY 2530 AGGACACTCAGCGAGTTCTGGATCACCTGTGTATGTCAACAGAGGATACCGTCTCCT 2589
 DB 2401 AGGACACTCAGCGAGTTCTGGATCACCTGTGTATGTCAACAGAGGATACCGTCTCCT 2460
 QY 2590 TGAAGAGGAAACTCTGTCACTCTCTCATGCTGTAGCTCATACACCAATTTCTCTTTGC 2649
 DB 2461 TGAAGAGGAAACTCTGTCACTCTCTCATGCTGTAGCTCATACACCAATTTCTCTTTGC 2520
 QY 2650 TTCAAGGTTTAAACTGGTTTTTGGATACCTCTATTAATTTCTCTGTCTCTCTCTGTT 2709
 DB 2521 TTCAAGGTTTAAACTGGTTTTTGGATACCTCTATTAATTTCTCTGTCTCTCTCTGTT 2580
 QY 2710 TATCTCTCCCTCCCTCCCTCCCTCTCTCTCCATCTCCATCTCTTTTGAATTTCCCTCAT 2769
 DB 2581 TATCTCTCCCTCCCTCCCTCCCTCTCTCTCCATCTCCATCTCTTTTGAATTTCCCTCAT 2640

QY	2770	CCCTCCATCTCAATCCCGTACTTACGCAACCCGCCCCCCCCCCAGCAAGCAGTGTCTCTGA	2829
Db	2641	CCCTCCATCTCAATCCCGTACTTACGCAACCCGCCCCCCCCCCAGCAAGCAGTGTCTCTGA	2700
QY	2830	GTATCATCATCACAAAGAGGAAACAAAGCGAAACACACAAACACAGCCTCAACTTACACTT	2889
Db	2701	GTATCATCATCACAAAGAGGAAACAAAGCGAAACACACAAACACAGCCTCAACTTACACTT	2760
QY	2890	GGTTACTCAAAAGAAAGAGTCAATGGTACTTGTCTAGCGTTTGGAGAGAGGAAACA	2949
Db	2761	GGTTACTCAAAAGAAAGAGTCAATGGTACTTGTCTAGCGTTTGGAGAGAGGAAACA	2820
QY	2950	GGAAACCCACAAACCAACCAATCAACAAACAAAGAAATAATCCCAATGAAGAAGATGT	3009
Db	2821	GGAAACCCACAAACCAACCAATCAACAAACAAAGAAATAATCCCAATGAAGAAGATGT	2880
QY	3010	ATTTTGTCTTTTTCATTTTGGTGATATAAGCCATCAATATTCAGCAAAATGATTCCTTTTC	3069
Db	2881	ATTTTGTCTTTTTCATTTTGGTGATATAAGCCATCAATATTCAGCAAAATGATTCCTTTTC	2940
QY	3070	TTTAAAGAAAAAATGTGGAGAAAGTAGAAATTTACCAAGGTGTGTGGCCCAAGGCGTT	3129
Db	2941	TTTAAAGAAAAAATGTGGAGAAAGTAGAAATTTACCAAGGTGTGTGGCCCAAGGCGTT	3000
QY	3130	AAATTCACAGATTTTTTAAACGAGAAAAACACACAGAAAGAGCTACCTCAGGTGTTTTTTA	3189
Db	3001	AAATTCACAGATTTTTTAAACGAGAAAAACACACAGAAAGAGCTACCTCAGGTGTTTTTTA	3060
QY	3190	CCTCAGACACCTTGCTTGTTGTTCCCTTAGAGATTTTGAAGCTGATAGTTGGAGCAT	3249
Db	3061	CCTCAGACACCTTGCTTGTTGTTCCCTTAGAGATTTTGAAGCTGATAGTTGGAGCAT	3120
QY	3250	TTTTTTATTTTTTAAATAAAATGAGTTGAAAAAAAATAAGATATCAACTGCCAGCGCTG	3309
Db	3121	TTTTTTATTTTTTAAATAAAATGAGTTGAAAAAAAATAAGATATCAACTGCCAGCGCTG	3180
QY	3310	GAGAAGGTGACAGTCCAAGTGTGCAACAGCTGTTCTGAAATGTCTCCGCTAGCCAGAA	3369
Db	3181	GAGAAGGTGACAGTCCAAGTGTGCAACAGCTGTTCTGAAATGTCTCCGCTAGCCAGAA	3240
QY	3370	CCNATATGGCTCTCTTTTGGACAAACCTTGAAATGTTTATTT	3412
Db	3241	CCNATATGGCTCTCTTTTGGACAAACCTTGAAATGTTTATTT	3283
RESULT 8			
AF117107			
LOCUS			
DEFINITION	Homo sapiens IGF-II mRNA-binding protein 2 (IMP-2) mRNA, complete cds.	2010 bp	linear
ACCESSION	AF117107		
VERSION	AF117107.1	GI:4191609	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
	Homo sapiens		
REFERENCE			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 2010)		
JOURNAL	Nielsen, J., Christensen, J., Lykke-Andersen, J., Johnsen, A.H.,		
MEDLINE	Wewer, U.M. and Nielsen, F.C.		
PUBMED	A family of insulin-like growth factor II mRNA-binding proteins		
AUTHORS	represses translation in late development		
REFERENCE	Mol. Cell. Biol. 19 (2), 1262-1270 (1999)		
	99108099		
	2 (bases 1 to 2010)		
	Nielsen, J., Christensen, J., Lykke-Andersen, J., Johnsen, A.H.,		
	Wewer, U.M. and Nielsen, F.C.		
TITLE	Direct Submission		
JOURNAL	Submitted (30-DEC-1998) Institute of Molecular Biology, University		
FEATURES	of Copenhagen, Soelvgade 83H, Copenhagen DK-1307, Denmark		
	Location/Qualifiers		
	1. .2010		

Qy	657	GGGGATCTCTGGTCC	CCACCCAGTTTGT	TGGTGCATCAT	CGGAAAGAGGAGGCTTGACCAT	7116			
Db	634	GGGGATCTCTGGTCC	CCACCCAGTTTGT	TGGTGCATCAT	CGGAAAGAGGAGGCTTGACCAT	693			
Qy	717	AAAGAAATCATCA	TAAGCAGACCCAGT	CCCCGGGTAGATAT	CCATAGAAAAAGAGAACTCTGG	776			
Db	694	AAAGAAATCATCA	TAAGCAGACCCAGT	CCCCGGGTAGATAT	CCATAGAAAAAGAGAACTCTGG	753			
Qy	777	AGCTGCAGAGAAG	CGCTGTACCAT	CCATGCCCCAGGAGGAGCTTCT	GAAAGCATGCCC	836			
Db	754	AGCTGCAGAGAAG	CGCTGTACCAT	CCATGCCCCAGGAGGAGCTTCT	GAAAGCATGCCC	813			
Qy	837	CATGATCTTTGA	AAATCATGCAGAAA	AGGAGGAGATGAG	CCAAACTAGCCGAAAGAGATTCC	896			
Db	814	CATGATCTTTGA	AAATCATGCAGAAA	AGGAGGAGATGAG	CCAAACTAGCCGAAAGAGATTCC	873			
Qy	897	TTTGAAATCTTGG	CAACATGCTTGGT	TGGAAGACTGAT	TGAAAAAGAGGCGAGAAA	956			
Db	874	TTTGAAATCTTGG	CAACATGCTTGGT	TGGAAGACTGAT	TGAAAAAGAGGCGAGAAA	933			
Qy	957	TTTGAAAGAAA	TTGAACATGA	AAACAGGACCAAGATAA	CAATCTCATCTTTTGAGAGATT	1016			
Db	934	TTTGAAAGAAA	TTGAACATGA	AAACAGGACCAAGATAA	CAATCTCATCTTTTGAGAGATT	993			
Qy	1017	GAGCATATACA	CCCCGGAAGAAC	CACTCACTGTGAAGGG	CACAGTTGAGGCCCTGTGCCAG	1076			
Db	994	GAGCATATACA	CCCCGGAAGAAC	CACTCACTGTGAAGGG	CACAGTTGAGGCCCTGTGCCAG	1053			
Qy	1077	TGCTGAGATAG	AGATTATGAAG	AGCTGCTGAGGCCCTTT	GAAAAATGATATGCTGGCTGT	1136			
Db	1054	TGCTGAGATAG	AGATTATGAAG	AGCTGCTGAGGCCCTTT	GAAAAATGATATGCTGGCTGT	1113			
Qy	1137	TAACCAACAG	CCAACTCTGAT	CCAGGGTTGAACCT	CAGCGCACTTGGCATCTTTTCAAC	1196			
Db	1114	TAACCAACAG	CCAACTCTGAT	CCAGGGTTGAACCT	CAGCGCACTTGGCATCTTTTCAAC	1173			
Qy	1197	AGGACTGTGCTG	TATCTTCCA	CAGAGAGGCCCCGGGAGCT	CCCCCGCTGCCCCCTTA	1256			
Db	1174	AGGACTGTGCTG	TATCTTCCA	CAGAGAGGCCCCGGGAGCT	CCCCCGCTGCCCCCTTA	1233			
Qy	1257	CAACCCCTTCA	CTACCTCCG	GATCTTCTCC	AGCTGTATCC	CCCCCATCACCAAGTTGG	1316		
Db	1234	CAACCCCTTCA	CTACCTCCG	GATCTTCTCC	AGCTGTATCC	CCCCCATCACCAAGTTGG	1293		
Qy	1317	CCGGTTCCG	CATCATCTTTAT	CCAGAGCAGGAGATTG	TGAATCTCTTCA	TCCCAAC	1376		
Db	1294	CCGGTTCCG	CATCATCTTTAT	CCAGAGCAGGAGATTG	TGAATCTCTTCA	TCCCAAC	1353		
Qy	1377	CAGGCTGTGGG	CCCATCAT	CGGAAAGAGGGG	CACATCAAA	CAGCTGCGGAGATT	1436		
Db	1354	CAGGCTGTGGG	CCCATCAT	CGGAAAGAGGGG	CACATCAAA	CAGCTGCGGAGATT	1413		
Qy	1437	CGCCGAG	CCCTCTAT	CAAGATTGCC	CTCGGAAAGGCC	CAGAGCTCAGCGAAAAGGATGGT	1496		
Db	1414	CGCCGAG	CCCTCTAT	CAAGATTGCC	CTCGGAAAGGCC	CAGAGCTCAGCGAAAAGGATGGT	1473		
Qy	1497	CATCATCA	CCGGG	CCACCGG	AGCTTCA	AGGCGGAGCGGATCTTTGGGAAACT	1556		
Db	1474	CATCATCA	CCGGG	CCACCGG	AGCTTCA	AGGCGGAGCGGATCTTTGGGAAACT	1533		
Qy	1557	GAAAGAGG	AAAACTTT	TTAA	CCCCCAAGAGAGTGA	AGCTCGAAGCGCATATCAGAT	1616		
Db	1534	GAAAGAGG	AAAACTTT	TTAA	CCCCCAAGAGAGTGA	AGCTCGAAGCGCATATCAGAT	1593		
Qy	1617	GCCCTCTT	CCACAG	CTGGCCG	GTGATG	TGCAAGAGTGGCA	AGCCGTGAA	CGAACTGCA	1676
Db	1594	GCCCTCTT	CCACAG	CTGGCCG	GTGATG	TGCAAGAGTGGCA	AGCCGTGAA	CGAACTGCA	1653
Qy	1677	GAATTTAA	CCAGTGCAG	AACTCAT	CTGCTCT	CTGTGACCA	AAACCGG	AGATGAA	1736
Db	1654	GAATTTAA	CCAGTGCAG	AACTCAT	CTGCTCT	CTGTGACCA	AAACCGG	AGATGAA	1713
Qy	1737	AGTGATCGT	CAGAA	TTATCGG	GCATCTTT	TGCTAGC	AGCTG	CA	1796

Db	1714	AGTGATGTCAGAAATTATCGGGCACTCTTTGTCGCCAGACTGCACAGCGCAAGATCAG	1773		
Qy	1797	GGAAATTTGTACAAACAGGTTGAAGCAGCAGCAGGAGCAAGAATAACCTTCAGGGAGTGCCTCACA	1856		
Db	1774	GGAAATTTGTACAAACAGGTTGAAGCAGCAGCAGGAGCAAGAATAACCTTCAGGGAGTGCCTCACA	1833		
Qy	1857	GGCGACCAAGTGAAGTCTCCACAGGCACACAGCAAAACAACGAGTGAATGTAGAGCCCTTCCA	1916		
Db	1834	GGCGACCAAGTGAAGTCTCCACAGGCACACAGCAAAACAACGAGTGAATGTAGAGCCCTTCCA	1893		
Qy	1917	ACACCTTGACAGATGAGACCAAAACGACGACGACGACGATCGGAGCAAAACCAAGAGACCATC	1976		
Db	1894	ACACCTTGACAGATGAGACCAAAACGACGACGACGACGATCGGAGCAAAACCAAGAGACCATC	1953		
Qy	1977	TGAGGAATGAGAAGTCTCGGAGGCGGCCAGGACTCTGCCGAGGCGCTTGAGAACCC	2033		
Db	1954	TGAGGAATGAGAAGTCTCGGAGGCGGCCAGGACTCTGCCGAGGCGCTTGAGAACCC	2010		
RESULT 9	AF057352	3667 bp	mrna	linear	PRI 23-MAY-1999
LOCUS	Homo sapiens hepatocellular carcinoma autoantigen (p62) mRNA,				
DEFINITION	complete cds.				
ACCESSION	AF057352				
VERSION	AF057352.1	GI:4883680			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
TITLE	1 (bases 1 to 3667)				
JOURNAL	Zhang, J.Y., Chan, E.K., Peng, X.X. and Tan, E.M.				
MEDLINE	A novel cytoplasmic protein with RNA-binding motifs is an				
PUBMED	autoantigen in human hepatocellular carcinoma				
REFERENCE	J. Exp. Med. 189 (7), 1101-1110 (1999)				
AUTHORS	2 (bases 1 to 3667)				
TITLE	Zhang, J.Y., Chan, E.K.L., Peng, X.X. and Tan, E.M.				
JOURNAL	Direct Submission				
FEATURES	Submitted (03-APR-1998) Molecular & Experimental Medicine, The				
Source	Scripps Research Institute, 10550 N. Torrey Pines Road, La Jolla,				
gene	CA 92037, USA				
CDS	Location/Qualifiers				
	1..3667				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	1..3667				
	/gene="p62"				
	436..2106				
	/gene="p62"				
	/note="RNA-binding protein"				
	/codon_start=1				
	/product="hepatocellular carcinoma autoantigen"				
	/protein_id="AAD31596.1"				
	/db_xref="GI:4883681"				
ORIGIN	/translation="MMNKLYIGNLSPAVTADDRLQLFGDRKULPLAQVLLKSGYAFVD				
	YPDQNWAIETLSGKVELHGMKIMEVDYSVKLSRKRQIRNIPPHIQWELDGLL				
	AOYGVENVEQNTDPTAVNVTVATREAAKIMKLSHQFPENYSFKISYIPDERV				
	SPSPQARQGDHSSREGHAPGTSQARQIDFPLRILVPTQFVGAIIKKEGLTIKN				
	IYKQTSRVDILHRKENSAAEKPTVIHATPEGTSEACRMILEIMKEADETKLAERIP				
	LKILANGVLVRLIGKEGNLKKIHEHTGKTIISLQDLISYNPRTIIVKGTVERAC				
	ASAEIEMKKLREAFENDMLAVNTHSGYFSSLYPHHPGFPFHHHSYPEOEIVNLFIP				
	TQAVGAIIGKKGAHIKQLARFAGASIKIAPAGPDVSRMVIITGPPEAKQKQRIIF				
	GKLKEBNFNFKEEVKLEAIRVPSATGRVIGKGTNVLQNLTSABYIVPRDQTP				
	DENEVIRIIGHFFASQTAQRKIREIVQVKQEQKYQGVASQSK"				
Query Match	53.0%	Score 1810;	DB 9;	Length 3667;	
Best Local Similarity	99.9%	Pred. No. 0;			

Matches 2100; Conservative 0; Mismatches 1; Indels 2; Gaps 2;			
QY	1270	ACCCATCCGGATACCTTCTCCAGCCTGTACCCCATCACAGATTTGGCCCGCTTCCCGCAT	1329
Db	1507	ACCCATCCGGATACCTTCTCCAGCCTGTACCCCATCACAGATTTGGCCCGCTTCCCGCAT	1566
QY	1330	CATCACTCTATCCAGAGCAGAGATGTGAATCTCTTATCCCAACCCAGGCTGTGGGC	1389
Db	1567	CATCACTCTATCCAGAGCAGAGATGTGAATCTCTTATCCCAACCCAGGCTGTGGGC	1626
QY	1390	GCATCATCGGGAAGAGGGGCACACATCAACACGCTGGCGAGATTCGCCGGAGCCTCT	1449
Db	1627	GCCATCATCGGGAAGAGGGGCACACATCAACACGCTGGCGAGATTCGCCGGAGCCTCT	1686
QY	1450	ATCAAGATTGCCCTCGGAAGGCCAGACGTCAGCGAAAGGATGTCTATCATCACCGGG	1509
Db	1687	ATCAAGATTGCCCTCGGAAGGCCAGACGTCAGCGAAAGGATGTCTATCATCACCGGG	1746
QY	1510	CCACCGGAAGCCAGTTCAAGGCCAGGACGGATCTTTGGGAACTGAAAGGGAAC	1569
Db	1747	CCACCGGAAGCCAGTTCAAGGCCAGGACGGATCTTTGGGAACTGAAAGGGAAC	1806
QY	1570	TTCTTTAAACCCAAAGAAAGTGAAGCTGGAAGCGCATATCAGAGTGCCTCTTCCACA	1629
Db	1807	TTCTTTAAACCCAAAGAAAGTGAAGCTGGAAGCGCATATCAGAGTGCCTCTTCCACA	1866
QY	1630	GCTGGCCGGTGATTTGGCAAGGTGGCAAGACCGTGAACGAACTGCAGAACTTAAACAGT	1689
Db	1867	GCTGGCCGGTGATTTGGCAAGGTGGCAAGACCGTGAACGAACTTAAACAGT	1926
QY	1690	GCAGAGTCATCGTGCCTCTGACCAAAACCGCCAGATGAAATCAGGAAGTGCATCAG	1749
Db	1927	GCAGAGTCATCGTGCCTCTGACCAAAACCGCCAGATGAAATCAGGAAGTGCATCAG	1986
QY	1750	ATTATCGGGCACTTTCTTTGCTAGCCAGACTGCACAGCGAAGATCAGGGAATTTACAA	1809
Db	1987	ATTATCGGGCACTTTCTTTGCTAGCCAGACTGCACAGCGAAGATCAGGGAATTTACAA	2046
QY	1810	CHAGTGAAGCAGCAGAGCAGAAATACCTCTAGGGAGTGCCTTCAAGCGCAGCAAGTGA	1869
Db	2047	CAGGTGAAGCAGCAGAGCAGAAATACCTCTAGGGAGTGCCTTCAAGCGCAGCAAGTGA	2106
QY	1870	GGCTCCACAGGCACAGCAAAACRAAGATGATGTAGCCCTTCCAAACACCTGCAGAA	1929
Db	2107	GGCTCCACAGGCACAGCAAAACRAAGATGATGTAGCCCTTCCAAACACCTGCAGAA	2166
QY	1930	TGAGACAAACGCGAGCCAGATCGGAGCAAAACAAAGACCATCTCAGGGAATCAGAA	1989
Db	2167	TGAGACAAACGCGAGCCAGATCGGAGCAAAACAAAGACCATCTCAGGGAATCAGAA	2226
QY	1990	GTCTGGGAGGCGGCCAGGGAATCTGCGGAGGCCCTTGAGAACCCCAAGGGCCGAGGAGG	2049
Db	2227	GTCTGGGAGGCGGCCAGGGAATCTGCGGAGGCCCTTGAGAACCCCAAGGGCCGAGGAGG	2286
QY	2050	GGGGGAAGGTCAGCCAGGTTTGCAGAACCCACGAGCCCGCCTCCGCCCCCGCAGGGC	2109
Db	2287	GGGGGAAGGTCAGCCAGGTTTGCAGAACCCACGAGCCCGCCTCCGCCCCCGCAGGGC	2346
QY	2110	TTCTGAGGCTTCAGCCATCCACTTCACCATCCACTCGGATCTCTCTGAACTCCCAAGA	2169
Db	2347	TTCTGAGGCTTCAGCCATCCACTTCACCATCCACTCGGATCTCTCTGAACTCCCAAGA	2406
QY	2170	CGTATCCCTTTTGTAGTTGAACATAAGGTGAACGTTTCAAGCCAGCAAAATGCAC	2229
Db	2407	CGTATCCCTTTTGTAGTTGAACATAAGGTGAACGTTTCAAGCCAGCAAAATGCAC	2466
QY	2230	ACCTTTTCTGTGGCAATCGTCTGTACATGTGTATCATATTTAGAAAGGGAAGATG	2289
Db	2467	ACCTTTTCTGTGGCAATCGTCTGTGTACATGTGTATCATATTTAGAAAGGGAAGATG	2526
QY	2290	TTAAGATATGTGCGCTGTGGTTTACACAGGTCCTGCGCGGTAAATATATTTAGAAAT	2349
Db	2527	TTAAGATATGTGCGCTGTGGTTTACACAGGTCCTGCGCGGTAAATATATTTAGAAAT	2586

QY	2350	AAATATACAAATAACTCAACTAACTCAATTTTTTAATCAATTAATTAATTTTTTCTTT	2409
Db	2587	AAATATACAAATAACTCAACTAACTCAATTTTTTAATCAATTAATTAATTTTTTCTTT	2646
QY	2410	TTAAAGAGAAACAGGCTTTTCTAGACTTTAAAGATAAAGCTTTTGGGAGGTCTCACGG	2469
Db	2647	TTAAAGAGAAACAGGCTTTTCTAGACTTTAAAGATAAAGCTTTTGGGAGGTCTCACGG	2706
QY	2470	TGTAGAGGAGCTTTGAGGCCACCGCACAAAATTCACCCAGAGGAAATCTCGTCGGA	2529
Db	2707	TGTAGAGGAGCTTTGAGGCCACCGCACAAAATTCACCCAGAGGAAATCTCGTCGGA	2766
QY	2530	AGGACATCACCGCAGTTCTGGATCACCTGTGTATGTCAACAGAGGGATACCGTCTCCT	2589
Db	2767	AGGACATCACCGCAGTTCTGGATCACCTGTGTATGTCAACAGAGGGATACCGTCTCCT	2826
QY	2590	TGAAGAGGAACTGTGTCACTCTCATGCTGTCTAGCTCATACACCCATTTCTCTTGC	2649
Db	2827	TGAAGAGGAACTGTGTCACTCTCATGCTGTCTAGCTCATACACCCATTTCTCTTGC	2886
QY	2650	TTTCAAGGTTTTAAACTGGTTTTTGGCATACCTGTATATAAATCTCTGTCTCTCTGTT	2709
Db	2887	TTTCAAGGTTTTAAACTGGTTTTTGGCATACCTGTATATAAATCTCTGTCTCTCTGTT	2946
QY	2710	TATCTCTCCCTCCCTCCCTCCCTTCTTCTCCATCTCCATTTCTTCAATTTCTCAT	2769
Db	2947	TATCTCTCCCTCCCTCCCTCCCTTCTTCTCCATCTCCATTTCTTCAATTTCTCAT	3006
QY	2770	CCCTCCATCTCAATCCCGTATCTAGCAACCCGCCCCCGCCCGGCAAGAGTCTCTCA	2829
Db	3007	CCCTCCATCTCAATCCCGTATCTAGCAACCCGCCCCCGCCCGGCAAGAGTCTCTCA	3065
QY	2830	GTATCACATFACACAAAAGGAAACAAAGCGAAACACAAAACAGCCTCAAATTACACTT	2889
Db	3066	GTATCACATFACACAAAAGGAAACAAAGCGAAACACAAAACAGCCTCAAATTACACTT	3125
QY	2890	GGTTACTCAAAAGAACAGAGTCAATGGTACTGTGCTTCTAGCGTTTTTGGAGGAGGAAA	2949
Db	3126	GGTTACTCAAAAGAACAGAGTCAATGGTACTGTGCTTCTAGCGTTTTTGGAGGAGGAAA	3185
QY	2950	GGAAACCCACCAACCAACCAATCAACCAACCAAGGAAAGGAAATTCACATGAAGATGT	3009
Db	3186	GGAAACCCACCAACCAACCAATCAACCAACCAAGGAAAGGAAATTCACATGAAGATGT	3245
QY	3010	ATTTTGTCTTTTGTGATTTTGGTGTATAGCCATCAATATTCAGCAAAATGATTCCTTTC	3069
Db	3246	ATTTTGTCTTTTGTGATTTTGGTGTATAGCCATCAATATTCAGCAAAATGATTCCTTTC	3305
QY	3070	TTTT-AAAAAAAATGTGGAGAAAGTAGAAATTTTCAAGGTTTGTGGCCCGAGGCGT	3128
Db	3306	TTTTAAAAAAAATGTGGAGAAAGTAGAAATTTTACCAAGGTTTGTGGCCCGAGGCGT	3365
QY	3129	TAAATTCAGAGATTTTAAACGAGAAACACACAGAGAGAGCTACCTCAGGTGTTTTT	3188
Db	3366	TAAATTCAGAGATTTTAAACGAGAAACACACAGAGAGAGCTACCTCAGGTGTTTTT	3425
QY	3189	ACCTCAGCACCTTGTCTTGTGTTTCCCTTAGAGATTTTGTAAAGCTGATGTGGAGCA	3248
Db	3426	ACCTCAGCACCTTGTCTTGTGTTTCCCTTAGAGATTTTGTAAAGCTGATGTGGAGCA	3485
QY	3249	TTTTTTTATTTTTTAATAAAATGAGTTGGAAAAAAAATAAGATATCACTGCCAGCCT	3308
Db	3486	TTTTTTTATTTTTTAATAAAATGAGTTGGAAAAAAAATAAGATATCACTGCCAGCCT	3545
QY	3309	GGAGAGGTGACAGTCCAGTGTCAACAGCTGTCTGAAATTTGTCTTCCTAGCCAGA	3368
Db	3546	GGAGAGGTGACAGTCCAGTGTCAACAGCTGTCTGAAATTTGTCTTCCTAGCCAGA	3605
QY	3369	ACC 3371	
Db	3606	ACC 3608	

```
RESULT 10
HSM806243      2640 bp      mRNA      linear      PRI 17-JUN-2003
LOCUS          Homo sapiens mRNA; cDNA DKFZp686i17155 (from clone DKFZp686i17155).
DEFINITION
ACCESSION      BX537583
VERSION        BX537583.1  GI:31873661
KEYWORDS
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
REFERENCE      1 (bases 1 to 2640)
AUTHORS        Bloecker,H., Boecher,M., Mewes,H.W., Weill,B., Amid,C., Osanger,A.,
                Fobo,G., Han,M. and Wiemann,S.
TITLE          Direct Submission
JOURNAL
COMMENT        Submitted (17-JUN-2003) MIPS, Ingolstaedter Landstr.1, D-85764
                Neuherberg, GERMANY
                Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
                Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
                sequenced by GBF (National Research Centre for Biotechnology Ltd.,
                Braunschweig/Germany) within the cDNA sequencing consortium of the
                German Genome Project.
                This clone (DKFZp686i17155) is available at the RZPD in Berlin.
                Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
                Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
                information about the clone and the sequencing project is available
                at http://mips.gsf.de/proj/cDNA/.
FEATURES
source         Location/Qualifiers
1..2640
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="3q27.2"
/clone="DKFZp686i17155"
/tissue_type="human colon endothel primary cell culture"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_SfiI; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
2619_polyA_site
ORIGIN
Query Match      43.3%; Score 1478; DB 9; Length 2640;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1596; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1774 CAGACTGCAGCGCAGATCAGGGAAATTGTACAAAGGTGAGCGAGCGAGCAGAAA 1833
DB 993 CAGACTGCAGCGCAGATCAGGGAAATTGTACAAAGGTGAGCGAGCGAGCAGAAA 1052
QY 1834 TACCCTCAGGGAGTGCCTCACAGCGCAGCAAGTGAGGCTCCCAACAGGCGACCAAC 1893
DB 1053 TACCCTCAGGGAGTGCCTCACAGCGCAGCAAGTGAGGCTCCCAACAGGCGACCAAC 1112
QY 1894 AAGCGATGAATGTAGCCCTTCCAAACACTGACAGAAATGAGACCAAAACGCGAGCCAGCAGA 1953
DB 1113 AAGCGATGAATGTAGCCCTTCCAAACACTGACAGAAATGAGACCAAAACGCGAGCCAGCAGA 1172
QY 1954 TCGGGAGCAACCAAGAACCATCTGAGGAATGAGAACTCTCGGAGGCGCGGAGGACTC 2013
DB 1173 TCGGGAGCAACCAAGAACCATCTGAGGAATGAGAACTCTCGGAGGCGCGGAGGACTC 1232
QY 2014 TGCGGAGGCGCTGAGAACCCAGGGCCGAGGAGGGCGGGGAGGTTCAGCCAGGTTTCG 2073
DB 1233 TGCGGAGGCGCTGAGAACCCAGGGCCGAGGAGGGCGGGGAGGTTCAGCCAGGTTTCG 1292
QY 2074 CAGAACCAAGCGCCCGCTCCGCCCCCGAGGCTTCTGCGAGGCTTCAGCCATCCACT 2133
DB 1293 CAGAACCAAGCGCCCGCTCCGCCCCCGAGGCTTCTGCGAGGCTTCAGCCATCCACT 1352
QY 2134 TCACCATCCACTCGGATCTCTCTGAACTCCCAAGCGCTATCCCTTTTAGTTGAACATA 2193
DB 1353 TCACCATCCACTCGGATCTCTCTGAACTCCCAAGCGCTATCCCTTTTAGTTGAACATA 1412
```

```
QY 2194 CATAGGTGAAGTGTCTTCAAGCCCAAGCAAAATGACACACCCCTTTTCTGTGCAATCGTC 2253
DB 1413 CATAGGTGAAGTGTCTTCAAGCCCAAGCAAAATGACACACCCCTTTTCTGTGCAATCGTC 1472
QY 2254 TCTGTACATGTGTACATATTAGAAAGGGAAGATGTTAAGATATGTGGCTGTGGGTGTTA 2313
DB 1473 TCTGTACATGTGTACATATTAGAAAGGGAAGATGTTAAGATATGTGGCTGTGGGTGTTA 1532
QY 2314 CACAGGTGCTGCAGCGGTAAATATATTTAGAAATATATATCAAAATACTCAACTAAC 2373
DB 1533 CACAGGTGCTGCAGCGGTAAATATATTTAGAAATATATATCAAAATACTCAACTAAC 1592
QY 2374 TCCAAATTTTAAATCAATTTATTTTCTTTTAAAGGAGAGCGGCTTTTCTA 2433
DB 1593 TCCAAATTTTAAATCAATTTATTTTCTTTTAAAGGAGAGCGGCTTTTCTA 1652
QY 2434 GACTTTAAAGATAAAGTCTTTGGGAGGTCTCAGGGTGTAGAGAGGAGCTTTGAGGGCAC 2493
DB 1653 GACTTTAAAGATAAAGTCTTTGGGAGGTCTCAGGGTGTAGAGAGGAGCTTTGAGGGCAC 1712
QY 2494 CCGCACAAAATTCACCCAGAGGGAATCTCGTGGGAAGGACACTCAGGCGAGTTCTGGAT 2553
DB 1713 CCGCACAAAATTCACCCAGAGGGAATCTCGTGGGAAGGACACTCAGGCGAGTTCTGGAT 1772
QY 2554 CACTGTGTATGTCAACAGAGGAGATACCGTCTCTTTGAAGAGGAAACTCTGTCACTCCT 2613
DB 1773 CACTGTGTATGTCAACAGAGGAGATACCGTCTCTTTGAAGAGGAAACTCTGTCACTCCT 1832
QY 2614 CATGCTGTCTAGCTCATACACCCATTTCTTTGCTTTCAGGTTTAAACCTGGTTTT 2673
DB 1833 CATGCTGTCTAGCTCATACACCCATTTCTTTGCTTTCAGGTTTAAACCTGGTTTT 1892
QY 2674 TGCATCTCTATATAATTTCTGTCTCTCTGTGTATCTCTCCCTCCCTCCCTCCCTCC 2733
DB 1893 TGCATCTCTATATAATTTCTGTCTCTCTGTGTATCTCTCCCTCCCTCCCTCCCTCC 1952
QY 2734 CTCTCTCTCATCTCTCAATTTCTTTGAATTTCTCTCATCCCTCCATCTCAATCCCGTACTA 2793
DB 1953 CTCTCTCTCATCTCTCAATTTCTTTGAATTTCTCTCATCCCTCCATCTCAATCCCGTACTA 2012
QY 2794 CGCACCCCCCCCCCCCCAGGCAAGCAGTGTCTCTGAGTATCATATCACACAAAGGAACA 2853
DB 2013 CGCACCCCCCCCCCCCCAGGCAAGCAGTGTCTCTGAGTATCATATCACACAAAGGAACA 2072
QY 2854 AAGCGGAAACACACAAACAGCCTCACTTACACTTGGTTACTCAAAAGAACAGAGTCA 2913
DB 2073 AAGCGGAAACACACAAACAGCCTCACTTACACTTGGTTACTCAAAAGAACAGAGTCA 2132
QY 2914 ATGCTACTTCTCTAGCGTTTGGAGAGGAAACAGGAAACCCACCAACCAATCA 2973
DB 2133 ATGCTACTTCTCTAGCGTTTGGAGAGGAAACAGGAAACCCACCAACCAATCA 2192
QY 2974 ACCAAACAAAGAAAAAATTCACAATGAAGAATGATTTTGTCTTTTTCATTTTGGTG 3033
DB 2193 ACCAAACAAAGAAAAAATTCACAATGAAGAATGATTTTGTCTTTTTCATTTTGGTG 2252
QY 3034 TATAAGCCATCAATATTCAGCAAAATGATTTCTTTTAAAGAAAAAATGAGGAGGA 3092
DB 2253 TATAAGCCATCAATATTCAGCAAAATGATTTCTTTTAAAGAAAAAATGAGGAGGA 2312
QY 3093 AAGTAGAAATTTTACCAGGTTTGTGGCCAGGCGGTAAATTCACAGATTTTTTTTAAACGA 3152
DB 2313 AAGTAGAAATTTTACCAGGTTTGTGGCCAGGCGGTAAATTCACAGATTTTTTTTAAACGA 2372
QY 3153 GAAAAACACACAGAGAGCTACTCAGGTGTTTTTACCTCAGGACCTTGTCTGTGTT 3212
DB 2373 GAAAAACACACAGAGAGAGCTACTCAGGTGTTTTTACCTCAGGACCTTGTCTGTGTT 2432
QY 3213 TCCTTTAGAGATTTTGTAAAGCTGATAGTTGGAGCATTTTTTTTATTTTTTAAATAAAAT 3272
DB 2433 TCCTTTAGAGATTTTGTAAAGCTGATAGTTGGAGCATTTTTTTTATTTTTTAAATAAAAT 2492
```


QY	2914	ATGGTACTTGTCTCTAGCGTTTGGAGAGGAAAAACAGGACCCACCAACCAATCA	2973
Db	26574	ATGGTACTTGTCTCTAGCGTTTGGAGAGGAAAAACAGGACCCACCAACCAATCA	26515
QY	2974	ACCAACCAAGAAAAATTCACAATGAAGAATGATTTTGTCTTTTGGCAATTTGGTG	3033
Db	26514	ACCAACCAAGAAAAATTCACAATGAAGAATGATTTTGTCTTTTGGCAATTTGGTG	26455
QY	3034	TATAAGCCATCAATATTCAGCAAAATGATTCCTTTCTTT-AAAAAATAAATGTGGAGA	3092
Db	26454	TATAAGCCATCAATATTCAGCAAAATGATTCCTTTCTTT-AAAAAATAAATAATGTGGAGA	26395
QY	3093	AGTGAATAATTTACCAAGGTTGTTCGCCCGAGGCGCTTAAATTCACAGATTTTTTAAACGA	3152
Db	26394	AGTGAATAATTTACCAAGGTTGTTCGCCCGAGGCGCTTAAATTCACAGATTTTTTAAACGA	26335
QY	3153	GAAAAACACACAGAGAAGCTACCTCAGGTGTTTTTACCTCAGCAGCTTGTCTTGTGTT	3212
Db	26334	GAAAAACACACAGAGAAGCTACCTCAGGTGTTTTTACCTCAGCAGCTTGTCTTGTGTT	26275
QY	3213	TCCTTTAGAGATTTTGTAAAGCTGATAGTTGGAGCATTTTTTTATTTTTTAAATAAAT	3272
Db	26274	TCCTTTAGAGATTTTGTAAAGCTGATAGTTGGAGCATTTTTTTATTTTTTAAATAAAT	26215
QY	3273	GAGTTGGAAAAAATAAGATATCAACTGCCAGCTCGAGAGGTCACAGTCCAGTGTG	3332
Db	26214	GAGTTGGAAAAAATAAGATATCAACTGCCAGCTCGAGAGGTCACAGTCCAGTGTG	26155
QY	3333	CAACAGCTGTTCTGAAATTTCTTCGCTAGCCCAAGAAC	3371
Db	26154	CAACAGCTGTTCTGAAATTTCTTCGCTAGCCCAAGAAC	26116
RESULT 12			
AC134942/c			
LOCUS	AC134942	208273 bp	DNA linear HTG 27-MAR-2003
DEFINITION	Homo sapiens chromosome 3 clone RP11-164H23, WORKING DRAFT		
ACCESSION	AC134942		
VERSION	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.		
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C., Alsbrooks, S.L., Amaratunge, H.C., Are, J.R., Ayele, M., Banks, T., Barbara, J., Benton, J., Binage, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, N., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, X., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseged, H., Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,		

Nguyen, N., Nickerson, E., Nwokenkwo, S., Oguh, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, I., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rivers, M., Rojas, A., Rojokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shooshtari, N., Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatik, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Umani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstock, G. and Gibbs, R.			
Direct Submission			
Unpublished			
2 (bases 1 to 208273)			
Worley, K.C.			
Direct Submission			
Submitted (03-OCT-2002)	Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
3 (bases 1 to 208273)			
Worley, K.C.			
Direct Submission			
Submitted (27-MAR-2003)	Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
On Mar 27, 2003	this sequence version replaced gi:27819431.		
-----	Genome Center		
Center:	Baylor College of Medicine		
Center code:	BCM		
Web site:	http://www.hgsc.bcm.tmc.edu/		
Drafting Center Code:	BCM		
Contact:	hgsc-help@bcm.tmc.edu		
-----	Project Information		
Center project name:	HDIO		
Center clone name:	RP11-164H23		
-----	Summary Statistics		
Sequencing vector:	Plasmid		
Chemistry:	Dye-terminator Big Dye; 100% of reads		
Assembly program:	Phrap; version 0.990329		
Consensus quality:	203806 bases at least Q40		
Consensus quality:	206019 bases at least Q30		
Consensus quality:	207337 bases at least Q20		
Estimated insert size:	200535; sum-of-contigs estimation		
Quality coverage:	10x in Q20 bases; sum-of-contigs estimation		

* NOTE:	Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).		
* NOTE:	This is a 'working draft' sequence. It currently consists of 21 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.		
*	1	1288:	contig of 1288 bp in length
*	1289	1388:	gap of unknown length
*	1389	2750:	contig of 1362 bp in length
*	2751	2850:	gap of unknown length
*	2851	4127:	contig of 1277 bp in length
*	4128	4227:	gap of unknown length
*	4228	5868:	contig of 1641 bp in length
*	5869	5968:	gap of unknown length
*	5969	7223:	contig of 1255 bp in length
*	7224	7323:	gap of unknown length
*	7324	8668:	contig of 1345 bp in length
*	8669	8768:	gap of unknown length
*	8769	10070:	contig of 1302 bp in length
*	10071	10171:	gap of unknown length
*	10171	11650:	contig of 1480 bp in length
*	11651	11750:	gap of unknown length

* 11751 13080: contig of 1330 bp in length
* 13081 13180: gap of unknown length
* 13181 14705: contig of 1525 bp in length
* 14706 14805: gap of unknown length
* 14806 15952: contig of 1147 bp in length
* 15953 16052: gap of unknown length
* 16053 17790: contig of 1738 bp in length
* 17791 17890: gap of unknown length
* 17891 20675: contig of 2785 bp in length
* 20676 20775: gap of unknown length
* 20776 24752: contig of 3977 bp in length
* 24753 24852: gap of unknown length
* 24853 35887: contig of 11035 bp in length
* 35888 35987: gap of unknown length
* 35988 51687: contig of 15700 bp in length
* 51688 51787: gap of unknown length
* 51788 67942: contig of 16155 bp in length
* 67943 68042: gap of unknown length
* 68043 99019: contig of 30977 bp in length
* 99020 99119: gap of unknown length
* 99120 132270: contig of 33151 bp in length
* 132271 132271: gap of unknown length
* 132371 16641: contig of 34271 bp in length
* 16642 16674: gap of unknown length
* 16674 208273: contig of 41532 bp in length.

FEATURES

Location/Qualifiers
1..208273
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"
/clone="RP11-164H23"

ORIGIN

Query Match 38.0%; Score 1296; DB 2; Length 208273;
Best local Similarity 99.8%; Pred. No. 0;
Matches 1596; Conservative 0; Mismatches 0; Indels 3; Gaps 2;

QY 1774 CAGACTGCAGCGCAAGATCAGGAAATTGTACAACAGGTGAAGCAGCAGGACAGAAA 1833
DB 82061 CAGACTGCAGCGCAAGATCAGGAAATTGTACAACAGGTGAAGCAGCAGGACAGAAA 82002
QY 1834 TACCTCAGGAGTGGCTCAGCGCAGCAAGTGAAGGCTCCACAGGCACACGCAAAAC 1893
DB 82001 TACCCTCAGGAGTGGCTCAGCGCAGCAAGTGAAGGCTCCACAGGCACACGCAAAAC 81942
QY 1894 AACGGATGATGTAGCCCTTCCACACTGACAGAAATGAGACCAAAACGAGCCAGCCAGA 1953
DB 81941 AACGGATGATGTAGCCCTTCCACACTGACAGAAATGAGACCAAAACGAGCCAGCCAGA 81882
QY 1954 TCGGGAGCAACCAAGAACCATCTGAGGAATGAGAAGTCTGCGGAGGCGCCAGGGACTC 2013
DB 81881 TCGGGAGCAACCAAGAACCATCTGAGGAATGAGAAGTCTGCGGAGGCGCCAGGGACTC 81822
QY 2014 TCGGAGGCCCTGAGAAACCCAGGGCCGAGGCGCGGGAGGAGTCAAGCAGGTTTGC 2073
DB 81821 TGCAGGCCCTGAGAAACCCAGGGCCGAGGCGCGGGAGGAGTCAAGCAGGTTTGC 81762
QY 2074 CAGAAACACCGAGCCCGCTCCGCCGCCAGGCTTCTGACAGGCTTCAGCCATCCACT 2133
DB 81761 CAGAAACACCGAGCCCGCTCCGCCGCCAGGCTTCTGACAGGCTTCAGCCATCCACT 81702
QY 2134 TCACCATCCACTCGGATCTCTCTGAACCTCCACGAGCTATCCCTTTTAGTTGAACATA 2193
DB 81701 TCACCATCCACTCGGATCTCTCTGAACCTCCACGAGCTATCCCTTTTAGTTGAACATA 81642
QY 2194 CATAGGTGAACGTGTTCAAGCCACGAAATGCAACCCCTTTTCTGTGGGCAATCGTC 2253
DB 81641 CATAGGTGAACGTGTTCAAGCCACGAAATGCAACCCCTTTTCTGTGGGCAATCGTC 81582
QY 2254 TCTGTATACGTGTGTACATATTAGAAGGGAAGATGTTAAGATATGCGCCTGTGGGTTA 2313
DB 81581 TCTGTATACGTGTGTACATATTAGAAGGGAAGATGTTAAGATATGCGCCTGTGGGTTA 81522

QY 2314 CACAGGTCCTGCGAGCGGTAATATATATTTTAGAAATATATATCAAAATCACTCAACTAAC 2373
DB 81521 CACAGGTCCTGCGAGCGGTAATATATTTTAGAAATATATATCAAAATCACTCAACTAAC 81462
QY 2374 TCCAAATTTAAATCAATATTAATTTTCTTTTAAAGAGAAAGCAGGCTTTTCTTA 2433
DB 81461 TCCAAATTTAAATCAATATTAATTTTCTTTTAAAGAGAAAGCAGGCTTTTCTTA 81402
QY 2434 GACTTTAAAGAAATAAGTCTTTGGAGGTCTCACGGTGTAGAGAGGAGCTTTTGAGGCCAC 2493
DB 81401 GACTTTAAAGAAATAAGTCTTTGGAGGTCTCACGGTGTAGAGAGGAGCTTTTGAGGCCAC 81342
QY 2494 CCGCAGAAATTCACCCAGAGGGAATCTCGTGGAGAGGACATCAAGGAGTCTGGAT 2553
DB 81341 CCGCAGAAATTCACCCAGAGGGAATCTCGTGGAGAGGACATCAAGGAGTCTGGAT 81282
QY 2554 CACTGTGTATGTCAACAGAGGAGATACCGTCTCTTTGAAGAGGAAATCTGTCACTCCT 2613
DB 81281 CACTGTGTATGTCAACAGAGGAGATACCGTCTCTTTGAAGAGGAAATCTGTCACTCCT 81222
QY 2614 CATGCTGTCTAGCTCATACACCAATTTCTTTGCTTCTCAGGTTTAAATCGGTTTT 2673
DB 81221 CATGCTGTCTAGCTCATACACCAATTTCTTTGCTTCTCAGGTTTAAATCGGTTTT 81162
QY 2674 TGCATCTGTATATTAATTTCTGTCTCTCTCTGTATCTCTCCCTCCCTCCCTCCC 2733
DB 81161 TGCATCTGTATATTAATTTCTGTCTCTCTCTGTATCTCTCCCTCCCTCCCTCCC 81102
QY 2734 CTCTCTCTCCATCTCCATTTCTTTGAATTTCTCATCCCTCCATCTCAATCCCTCATCTA 2793
DB 81101 CTCTCTCTCCATCTCCATTTCTTTGAATTTCTCATCCCTCCATCTCAATCCCTCATCTA 81042
QY 2794 CGCACCCCCCCCCCGAGCAAGAGCTCTGTAGTATCAATCAATCAACAAAGGAACA 2853
DB 81041 CGCA--CCCCCCCCCGAGCAAGAGCTCTGTAGTATCAATCAATCAACAAAGGAACA 80984
QY 2854 AAAGCGAAACACAAACAGAGCTCACTTACACTTGGTTACTCAAAGAACAGAGTCA 2913
DB 80983 AAAGCGAAACACAAACAGAGCTCACTTACACTTGGTTACTCAAAGAACAGAGTCA 80924
QY 2914 ATGTACTTGTCTAGCGTTTGGAGAGAAACAGAAACCCACCAACCAACCAATCA 2973
DB 80923 ATGTACTTGTCTAGCGTTTGGAGAGAAACAGAAACCCACCAACCAACCAATCA 80864
QY 2974 ACCAAACAAAGAAAAATTCACAAATGAAGAATGATTTTGTCTTTTGTCAATTTGGTG 3033
DB 80863 ACCAAACAAAGAAAAATTCACAAATGAAGAATGATTTTGTCTTTTGTCAATTTGGTG 80804
QY 3034 TATAAGCATCAATATTCAGCAAAATGATTCCTTTCTTT--AAAAAATAATGTGGAGGA 3092
DB 80803 TATAAGCATCAATATTCAGCAAAATGATTCCTTTCTTTAATAAATAATGTGGAGGA 80744
QY 3093 AAGTAGAAATTTACCAAGGTTTGGCCCGAGGCGTTAAATTCACAGATTTTTTAAACGA 3152
DB 80743 AAGTAGAAATTTACCAAGGTTTGGCCCGAGGCGTTAAATTCACAGATTTTTTAAACGA 80684
QY 3153 GAAAAACACAGAGAGAGTACTCAGGTGTTTTTACCTCAGACCTTGTCTTGTGTT 3212
DB 80683 GAAAAACACAGAGAGAGTACTCAGGTGTTTTTACCTCAGACCTTGTCTTGTGTT 80624
QY 3213 TCCCTTAGAGATTTGTAAAGCTGATAGTTGGAGCATTTTTTTTATTTTAAATAAAT 3272
DB 80623 TCCCTTAGAGATTTGTAAAGCTGATAGTTGGAGCATTTTTTTTATTTTAAATAAAT 80564
QY 3273 GAGTTGGAAAAAATAAGATATCACTGCCAGCTGGAGAGGTGACAGTCCAAAGTGT 3332
DB 80563 GAGTTGGAAAAAATAAGATATCACTGCCAGCTGGAGAGGTGACAGTCCAAAGTGT 80504
QY 3333 CAACAGCTGTCTGAATTTGTCTTCGCTAGCCCAAGAAC 3371
DB 80503 CAACAGCTGTCTGAATTTGTCTTCGCTAGCCCAAGAAC 80465


```
RESULT 13
AC126373/c
LOCUS
DEFINITION Homo sapiens chromosome 18 clone CTD-2307I13 map 18, LOW-PASS
SEQUENCE SAMPLING.
AC126373
ACCESSION AC126373.1 GI:21699324
VERSION HTG; HTGS PHASE0.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 50320)
AUTHORS Birren,B., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 18, clone CTD-2307I13
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 50320)
AUTHORS Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Illiev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mienga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
DIRECT SUBMISSION
TITLE Submitted (05-JUL-2002) Whitehead Institute/MIT Center for Genome
JOURNAL Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L27661
Center clone name: 2307_I_13
-----
* NOTE: This record contains 59 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1 766: contig of 766 bp in length
* 767 866: gap of 100 bp
* 867 1574: contig of 708 bp in length
* 1575 1674: gap of 100 bp
* 1675 2412: contig of 738 bp in length
* 2413 2512: gap of 100 bp
* 2513 3283: contig of 771 bp in length
* 3284 3384: gap of 100 bp
* 3384 4124: contig of 741 bp in length
* 4125 4224: gap of 100 bp
* 4225 4938: contig of 714 bp in length
* 4939 5038: gap of 100 bp
* 5039 5820: contig of 782 bp in length
* 5821 5920: gap of 100 bp
* 5921 6685: contig of 765 bp in length
* 6686 6785: gap of 100 bp
* 6786 7580: contig of 795 bp in length
* 7581 7680: gap of 100 bp
* 7681 8432: contig of 752 bp in length
* 8433 8532: gap of 100 bp
* 8533 9281: contig of 749 bp in length
* 9282 9381: gap of 100 bp
* 9382 10136: contig of 755 bp in length
* 10137 10236: gap of 100 bp
* 10237 10967: contig of 731 bp in length
* 10968 11067: gap of 100 bp
* 11068 11787: contig of 720 bp in length
* 11788 11887: gap of 100 bp
* 11888 12655: contig of 768 bp in length
* 12656 12755: gap of 100 bp
* 12756 13497: contig of 742 bp in length
* 13498 13597: gap of 100 bp
* 13598 14365: contig of 768 bp in length
* 14366 14465: gap of 100 bp
* 14466 15203: contig of 738 bp in length
* 15204 15303: gap of 100 bp
* 15304 16010: contig of 707 bp in length
* 16011 16110: gap of 100 bp
* 16111 16872: contig of 762 bp in length
* 16873 16972: gap of 100 bp
* 16973 17715: contig of 743 bp in length
* 17716 17815: gap of 100 bp
* 17816 18587: contig of 772 bp in length
* 18588 18687: gap of 100 bp
* 18688 19464: contig of 777 bp in length
* 19465 19564: gap of 100 bp
* 19565 20303: contig of 739 bp in length
* 20304 20403: gap of 100 bp
* 20404 21135: contig of 732 bp in length
* 21136 21235: gap of 100 bp
* 21236 21999: contig of 764 bp in length
* 22000 22099: gap of 100 bp
* 22099 22846: contig of 747 bp in length
* 22847 22946: gap of 100 bp
* 22947 23709: contig of 763 bp in length
* 23710 23809: gap of 100 bp
* 23810 24525: contig of 716 bp in length
* 24526 24625: gap of 100 bp
* 24626 25371: contig of 746 bp in length
* 25372 25471: gap of 100 bp
* 25472 26239: contig of 768 bp in length
* 26240 26339: gap of 100 bp
* 26340 27090: contig of 751 bp in length
* 27091 27190: gap of 100 bp
* 27191 27946: contig of 756 bp in length
* 27947 28046: gap of 100 bp
* 28047 28088: contig of 762 bp in length
* 28089 28908: gap of 100 bp
* 28909 29669: contig of 761 bp in length
* 29670 29769: gap of 100 bp
* 29770 30550: contig of 781 bp in length
* 30551 30650: gap of 100 bp
* 30651 31423: contig of 773 bp in length
* 31424 31523: gap of 100 bp
* 31524 32279: contig of 756 bp in length
* 32280 32379: gap of 100 bp
* 32380 33129: contig of 750 bp in length
* 33130 33229: gap of 100 bp
* 33230 33984: contig of 755 bp in length
* 33985 34084: gap of 100 bp
* 34085 34828: contig of 744 bp in length
* 34829 35665: contig of 737 bp in length
* 35666 35765: gap of 100 bp
```

* 35766 36535: contig of 770 bp in length
* 36536 gap of 100 bp
* 36636 37412: contig of 777 bp in length
* 37413 37512: gap of 100 bp
* 37513 38284: contig of 772 bp in length
* 38285 38384: gap of 100 bp
* 38385 39173: contig of 789 bp in length
* 39174 39273: gap of 100 bp
* 39274 40051: contig of 778 bp in length
* 40052 40151: gap of 100 bp
* 40152 40900: contig of 749 bp in length
* 40901 41000: gap of 100 bp
* 41001 41778: contig of 778 bp in length
* 41779 41878: gap of 100 bp
* 41879 42645: contig of 767 bp in length
* 42646 42745: gap of 100 bp
* 42746 43454: contig of 709 bp in length
* 43455 43554: gap of 100 bp
* 43555 44314: contig of 760 bp in length
* 44315 44414: gap of 100 bp
* 44415 45158: contig of 744 bp in length
* 45159 45258: gap of 100 bp
* 45259 46040: contig of 782 bp in length
* 46041 46140: gap of 100 bp
* 46141 46894: contig of 754 bp in length
* 46895 46994: gap of 100 bp
* 46995 47754: contig of 760 bp in length
* 47755 47854: gap of 100 bp
* 47855 48596: contig of 742 bp in length
* 48597 48696: gap of 100 bp
* 48697 49446: contig of 750 bp in length
* 49447 49546: gap of 100 bp
* 49547 50320: contig of 774 bp in length.

FEATURES
 source
 1..50320
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="18"
 /map="18"
 /clone="CTD-2307f13"
 /clone_lib="CITD1 Human BAC"

ORIGIN
Query Match 17.2%; Score 586; DB 2; Length 50320;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 586; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2205 GTGTTCAAAGCCAAAGTGCACACCCCTTTTCTGTGGCAATCGTCTGTACATGT 2264
Db 10940 GTGTTCAAAGCCAAAGTGCACACCCCTTTTCTGTGGCAATCGTCTGTACATGT 10881

QY 2265 GTGTACATATAGAAGGAGATGTTAAGATATGCGCTGTGGTTACACAGGGTGC 2324
Db 10880 GTGTACATATAGAAGGAGATGTTAAGATATGCGCTGTGGTTACACAGGGTGC 10821

QY 2325 TGCAGCGGTATATATTTTGAATAATATATCAATAACTCAACTAACTCCAAATTTTA 2384
Db 10820 TGCAGCGGTATATATTTTGAATAATATATCAATAACTCAACTAACTCCAAATTTTA 10761

QY 2385 ATCAATATTAATTTTTTTTTCTTTTAAAGAGAAACAGCGCTTTCTAGACATTTAAAGA 2444
Db 10760 ATCAATATTAATTTTTTTTTCTTTTAAAGAGAAACAGCGCTTTCTAGACATTTAAAGA 10701

QY 2445 ATAAAGTCTTTGGAGTCTCAGGTGTAGAGAGGCTTTGAGGCCACCGCACAAAT 2504
Db 10700 ATAAAGTCTTTGGAGTCTCAGGTGTAGAGAGGCTTTGAGGCCACCGCACAAAT 10641

QY 2505 TCACCCAGAGGGAAATCTGTCGGAAGGACACCTCAGCGCATGTTCTGGATCACCTGTGTAT 2564
Db 10640 TCACCCAGAGGGAAATCTGTCGGAAGGACACCTCAGCGCATGTTCTGGATCACCTGTGTAT 10581

QY 2565 GTCACAGAGAGGATACCGTCTCTTGAAGAGAAACTCTGTCACTCTCATGCTGTCT 2624

Db 10580 GTCAACAGAGGGATACCGTCTCTTGAAGAGAAACTCTGTCACTCTCTCATGCTGTCT 10521

QY 2625 AGCTCATACACCCATTTCTCTTTTGTCTTTCACAGAGTTTAAAGTGGTTTTCATACACTGCT 2684

Db 10520 AGCTCATACACCCATTTCTCTTTTGTCTTTCACAGAGTTTAAAGTGGTTTTCATACACTGCT 10461

QY 2685 ATATAATCTCTGTCTCTCTCTGTTTATCTCTCCCTCCCTCCCTCCCTCCCTCTCTTCTTCCCA 2744

Db 10460 ATATAATCTCTGTCTCTCTCTGTTTATCTCTCCCTCCCTCCCTCCCTCTCTTCTTCCCA 10401

QY 2745 TCTCCATCTTTTGAATTTCTCATCTCCCTCATCTCAATCCCGTAT 2790

Db 10400 TCTCCATCTTTTGAATTTCTCATCTCCCTCATCTCAATCCCGTAT 10355

RESULT 14
BD221443

LOCUS BD221443 710 bp DNA linear PAT 17-JUL-2003

DEFINITION Human gene and gene expression product V.

ACCESSION BD221443

VERSION BD221443.1 GI:33031213

KEYWORDS JP 2002534055-A/2556.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 710)

AUTHORS Williams,J.L., Escobedo,J., Innis,M.A., Garcia,P.D., Klinger,J.S., Reinhard,C., Giese,K., Randazzo,F., Kennedy,G.C., Pot,D., Kassam,A., Lamson,G., Drmanac,R., Crkvenjakov,R., Dickson,M., Drmanac,S., Labat,I., Leshkowitz,D., Kita,D., Garcia,V., Jones,L.W. and Crain,B.S.

TITLE Human gene and gene expression product V

JOURNAL Patent: JP 2002534055-A 2556 15-OCT-2002; CHIRON CORP,HYSEQ INC

COMMENT OS Homo sapiens (human)
PN JP 2002534055-A/2556
PD 15-OCT-2002
PF 13-MAY-1999 JP 2000548466
PR 14-MAY-1998 US 60/085426,15-MAY-1998 US 60/085537 PR 15-MAY-1998 US 60/085696,21-OCT-1998 US 60/105234 PR 27-OCT-1998 US 60/105877

PI LOUIS T WILLIAMS,JAIME ESCOBEDO,MICHAEL A INNIS,PABLO PI DOMINGUEZ GARCIA,
PI JULIE SUDDUTH KLINGER,CHRISTOPH REINHARD,KLAUSE GIESE,FILIPPO RANDAZZO,
PI GIULIA C KENNEDY,DAVID POT,ALTAF KASSAM,GEORGE LAMSON,RADOJE DRMANAC,
PI RADOMIR CRKVENJAKOV,MARK DICKSON,SNEZANA DRMANAC,IVAN LABAT, DENA LESHKOWITZ,DAVID KITA,VERONICA GARCIA,LEE WILLIAM JONES, BIRGIT STACHE CRAIN
PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21 PC C12N5/00,C12Q1/68,
PC C12N15/00,C12N5/00
n = A,T,C or G
FH Key Location/Qualifiers
FT misc location (i)..(710).

FEATURES
 source
 1..710
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

ORIGIN
Query Match 16.6%; Score 567; DB 6; Length 710;
Best Local Similarity 100.0%; Pred. No. 2.5e-309;
Matches 567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 534 CCGGATGAAGAGTGTGAGCTCCCTTCGCCCTCAGCGAGCCAGCGTGGGACCACTC 593

Db 25 CCGGATGAAGAGTGTGAGCTCCCTTCGCCCTCAGCGAGCCAGCGTGGGACCACTC 84

```
QY 594 TTCCCGGAGCAGGCGCCCTGGGGGCACTTCTCAGGCGAGACAGATTGATTTC 653
D 85 TTCCCGGAGCAGGCGCCCTGGGGGCACTTCTCAGGCGAGACAGATTGATTTC 144
QY 654 GCTGCGGATCTCTGCTCCACCCAGTTTGTGTGTCATCATCGGAAAGGAGGCTTGAC 713
D 145 GCTGCGGATCTCTGCTCCACCCAGTTTGTGTGTCATCATCGGAAAGGAGGCTTGAC 204
QY 714 CATAAAGAACATCACTAAGCAGACCCAGTCCCGGGTAGATATCCATAGAAAAGAGAACTC 773
D 205 CATAAAGAACATCACTAAGCAGACCCAGTCCCGGGTAGATATCCATAGAAAAGAGAACTC 264
QY 774 TGGAGTGCAGAGAAGCCTGTGACCATCATGTCACCCAGAGGGGACCTTCTGAAGCATG 833
D 265 TGGAGTGCAGAGAAGCCTGTGACCATCATGTCACCCAGAGGGGACCTTCTGAAGCATG 324
QY 834 CGCGATGATCTTGAATCATGCAGAAAGAGGAGATGAGACCAACTAGCCGAGAGAT 893
D 325 CGCGATGATCTTGAATCATGCAGAAAGAGGAGATGAGACCAACTAGCCGAGAGAT 384
QY 894 TCCTCTGAAAATCTTGGGCACAACTGCTTGGTGAAGACTGATTTGAAAAGAGGCGAG 953
D 385 TCCTCTGAAAATCTTGGGCACAACTGCTTGGTGAAGACTGATTTGAAAAGAGGCGAG 444
QY 954 AAATTTGAAGAAATTTGAATCAATGAACAGGAGCAGATGAACAACTCTCATCTTTGCAGGA 1013
D 445 AAATTTGAAGAAATTTGAATCAATGAACAGGAGCAGATGAACAACTCTCATCTTTGCAGGA 504
QY 1014 TTGAGCATATACAACCCGGAAGAACCATCTGTTGAAGGACAGTTTGAGGCTTGTGC 1073
D 505 TTGAGCATATACAACCCGGAAGAACCATCTGTTGAAGGACAGTTTGAGGCTTGTGC 564
QY 1074 CAGTGTGATAGATAGATTATGAAGAA 1100
D 565 CAGTGTGATAGATAGATTATGAAGAA 591

RESULT 15
AC126373
LOCUS
DEFINITION Homo sapiens chromosome 18 clone CTD-2307113 map 18, LOW-PASS
SEQUENCE SAMPLING.
AC126373
VERSION AC126373.1 GI:21699324
KEYWORDS HTG; HTGS_PHASE0.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 50320)
Birren,B., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 18, clone CTD-2307113
Unpublished
REFERENCE 2 (bases 1 to 50320)
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barn,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B.,
Camara,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-pierre,N., Hagos,B.,
Horton,L., Hulne,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karacas,A., Kells,C., Lander,T., Levine,R., Lindblad-Toh,K.,
Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
TITLE
```

JOURNAL

COMMENT

Submitted (05-JUL-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/BM/RepeatMasker.html>
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L27661
Center clone name: 2307_I_13

* NOTE: This record contains 59 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 766: contig of 766 bp in length
* 767 866: gap of 100 bp
* 867 1574: contig of 708 bp in length
* 1575 1674: gap of 100 bp
* 1675 2412: contig of 738 bp in length
* 2413 2512: gap of 100 bp
* 2513 3283: contig of 771 bp in length
* 3284 3383: gap of 100 bp
* 3384 4124: contig of 741 bp in length
* 4125 4224: gap of 100 bp
* 4225 4938: contig of 714 bp in length
* 4939 5038: gap of 100 bp
* 5039 5820: contig of 782 bp in length
* 5821 5920: gap of 100 bp
* 5921 6685: contig of 765 bp in length
* 6686 6785: gap of 100 bp
* 6786 7580: contig of 795 bp in length
* 7581 7680: gap of 100 bp
* 7681 8432: contig of 752 bp in length
* 8433 8532: gap of 100 bp
* 8533 9281: contig of 749 bp in length
* 9282 9381: gap of 100 bp
* 9382 10136: contig of 755 bp in length
* 10137 10236: gap of 100 bp
* 10237 10967: contig of 731 bp in length
* 10968 11067: gap of 100 bp
* 11068 11787: contig of 720 bp in length
* 11788 11887: gap of 100 bp
* 11888 12655: contig of 768 bp in length
* 12656 12755: gap of 100 bp
* 12756 13497: contig of 742 bp in length
* 13498 13597: gap of 100 bp
* 13598 14365: contig of 768 bp in length
* 14366 14465: gap of 100 bp
* 14466 15203: contig of 738 bp in length
* 15204 15303: gap of 100 bp
* 15304 16010: contig of 707 bp in length
* 16011 16110: gap of 100 bp
* 16111 16872: contig of 762 bp in length
* 16873 16972: gap of 100 bp
* 16973 17715: contig of 743 bp in length
* 17716 17815: gap of 100 bp
* 17816 18587: contig of 772 bp in length
* 18588 18687: gap of 100 bp
* 18688 19464: contig of 777 bp in length
* 19465 19564: gap of 100 bp
* 19565 20303: contig of 739 bp in length
* 20304 20403: gap of 100 bp

20404 21135: contig of 732 bp in length
21136 gap of 100 bp
21235: contig of 764 bp in length
21236 gap of 100 bp
22000 22099: contig of 747 bp in length
22100 22846: contig of 747 bp in length
22847 gap of 100 bp
22947 23709: contig of 763 bp in length
23710 23809: gap of 100 bp
23810 24525: contig of 716 bp in length
24526 gap of 100 bp
24626 25371: contig of 746 bp in length
25372 25471: gap of 100 bp
25472 26239: contig of 768 bp in length
26240 26339: gap of 100 bp
26340 27090: contig of 751 bp in length
27091 27190: gap of 100 bp
27191 27946: contig of 756 bp in length
27947 gap of 100 bp
28047 28808: contig of 762 bp in length
28809 28908: gap of 100 bp
28909 29669: contig of 761 bp in length
29670 29769: gap of 100 bp
29770 30550: contig of 781 bp in length
30551 30650: gap of 100 bp
30651 31423: contig of 773 bp in length
31424 31524: gap of 100 bp
31524 32279: contig of 756 bp in length
32280 32379: gap of 100 bp
32380 33129: contig of 750 bp in length
33130 33229: gap of 100 bp
33230 33984: contig of 755 bp in length
33985 34084: gap of 100 bp
34085 34828: contig of 744 bp in length
34829 34928: gap of 100 bp
34929 35665: contig of 737 bp in length
35666 35765: gap of 100 bp
35766 36535: contig of 770 bp in length
36536 36635: gap of 100 bp
36636 37412: contig of 777 bp in length
37413 37512: gap of 100 bp
37513 38284: contig of 772 bp in length
38285 38384: gap of 100 bp
38385 39173: contig of 789 bp in length
39174 39274: gap of 100 bp
39274 40051: contig of 778 bp in length
40052 40151: gap of 100 bp
40152 40900: contig of 743 bp in length
40901 41000: gap of 100 bp
41001 41778: contig of 778 bp in length
41779 41878: gap of 100 bp
41879 42645: contig of 767 bp in length
42646 42745: gap of 100 bp
42746 43554: contig of 709 bp in length
43555 44314: gap of 100 bp
44315 44414: contig of 760 bp in length
44415 45158: contig of 744 bp in length
45159 45258: gap of 100 bp
45259 46040: contig of 782 bp in length
46041 46140: gap of 100 bp
46141 46894: contig of 754 bp in length
46895 46994: gap of 100 bp
46995 47754: contig of 760 bp in length
47755 47854: gap of 100 bp
47855 48596: contig of 742 bp in length
48597 gap of 100 bp
48697 49446: contig of 750 bp in length
49447 49546: gap of 100 bp
49547 50320: contig of 774 bp in length.

FEATURES
Source

Location/Qualifiers
1. .50320
/organism="Homo sapiens"
/mol_type="genomic DNA"

/db_xref="taxon:9606"
/chromosome="18"
/map="18"
/clone="CTD-2307113"
/clone_lib="CITd1 Human BAC"

ORIGIN

Query Match
Best Local Similarity 16.6%; Score 565; DB 2; Length 50320;
Matches 615; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1917 ACACCTGACAGAAATGAGACCAACGCGAGCCAGATCGGAGCAACCAAGACCAATC 1976
Db 37566 ACACCTGACAGAAATGAGACCAACGCGAGCCAGATCGGAGCAACCAAGACCAATC 37625

QY 1977 TGAGGAATGAGAACTCTGCGAGGCGGCGAGGACTCTGCGAGGCGGCTTGAGAACCCGAG 2036
Db 37626 TGAGGAATGAGAACTCTGCGAGGCGGCGAGGACTCTGCGAGGCGGCTTGAGAACCCGAG 37685

QY 2037 GGGCGGAGGAGGCGGCGGAGGTCAGCCAGGTTTGCAGAACCAACCGAGCCGCGCTCC 2096
Db 37686 GGGCGGAGGAGGCGGCGGAGGTCAGCCAGGTTTGCAGAACCAACCGAGCCGCGCTCC 37745

QY 2097 CGCCCCCGAGGCTTCTGCGAGGCTTTCAGCCATCCACTTCCACTCCACTCGGATCTCTCC 2156
Db 37746 CGCCCCCGAGGCTTCTGCGAGGCTTTCAGCCATCCACTTCCACTCGGATCTCTCC 37805

QY 2157 TGAACCTCCACGACGCTATCCCTTTAGTTGAACTAAACATAGGTGAACGTTTCAAGGCC 2216
Db 37806 TGAACCTCCACGACGCTATCCCTTTAGTTGAACTAAACATAGGTGAACGTTTCAAGGCC 37865

QY 2217 AAGCAAAATGCACACCCCTTTTCTGTCGCAAAATCGTCTGTACATGTGTACATATTA 2276
Db 37866 AAGCAAAATGCACACCCCTTTTCTGTCGCAAAATCGTCTGTACATGTGTACATATTA 37925

QY 2277 GAAAGGGAAGATGTTAAGATATGTGGCTGTGGGTTACACAGGCTGCTGAGCGGTAAT 2336
Db 37926 GAAAGGGAAGATGTTAAGATATGTGGCTGTGGGTTACACAGGCTGCTGAGCGGTAAT 37985

QY 2337 ATATTTAGAAATATATATCAATAACTCAACTCAACTTCAACTTCAACTTCAACTTCAACTTCA 2396
Db 37986 ATATTTAGAAATATATATCAATAACTCAACTCAACTTCAACTTCAACTTCAACTTCAACTT 38045

QY 2397 TTTTCTTTCTTTTAAAGAGAAAGCAGGCTTTTCTAGACTTTTAAAGAAATAAAGTCTTTG 2456
Db 38046 TTTTCTTTCTTTTAAAGAGAAAGCAGGCTTTTCTAGACTTTTAAAGAAATAAAGTCTTTG 38105

QY 2457 GGAGGTCACGGTGTAGAGAGGAGCTTTTGAGGCCACCCGACAAAATTTCAACCCAGAGGG 2516
Db 38106 GGAGGTCACGGTGTAGAGAGGAGCTTTTGAGGCCACCCGACAAAATTTCAACCCAGAGGG 38165

QY 2517 AAATCTCGTGGGAGG 2532
Db 38166 AAATCTCGTGGGAGG 38181

Search completed: July 21, 2004, 16:45:49
Job time : 12978 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 21, 2004, 12:23:17 ; Search time 1236 Seconds

Title: US-09-270-437D-6
Perfect score: 3412
Sequence: 1 ggcagcggaggcggagga.....aaccttgaaatgtttattt 3412
11727.226 Million cell updates/sec

Scoring table: OLIGO_NUC
Gapop_60.0 , Gapext 60.0

Searched: 3373863 seqs, 2124099041 residues
Word size : 0

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB	ID	Description
1	3412	100.0	3412	3	AAZ36152	AAZ36152 DNA encod
2	2953	86.5	3694	4	AAS26150	Human CDN
3	2953	86.5	3694	7	ABX73491	Human nov
4	2143	62.8	3283	3	AAZ36154	AAZ36154 An altern
5	1977	57.9	2010	5	AAS70981	DNA encod
6	1810	53.0	3667	5	AAS70982	DNA encod
7	1810	53.0	3667	6	ABS76442	CDNA enco
8	1496	43.8	2290	4	AAZ26153	Human CDN
9	1496	43.8	2290	7	ABX73494	Human nov
10	567	16.6	710	3	AAZ02565	Human col
11	401	11.8	1186	4	AAS26570	Human CDN
12	401	11.8	1186	7	ABX73911	Human nov
13	389	11.4	822	4	AAZ26151	Human CDN
14	389	11.4	822	7	ABX73492	Human nov
15	347	10.2	364	3	AAC03267	Human sec
16	290	8.5	424	3	AAH31011	Human col
17	285	8.4	300	3	AAZ01526	Human col
18	264	7.7	264	4	AAZ24943	Human col
19	264	7.7	364	5	AAH83588	Human ova
20	263	7.7	263	4	AAZ25126	Human ova
21	263	7.7	263	4	AAZ25356	Human ova
22	263	7.7	378	5	AAH83772	Human ova
23	222	6.5	222	4	AAS58000	CDNA #676

24	222	6.5	222	4	AAS58609	CDNA #128
25	218	6.4	372	9	ADD71082	Human IGF
26	167	4.9	218	4	AAS58739	CDNA #141
27	156	4.6	306	2	AAV87489	EST clone
28	143	4.2	402	8	ACH18272	Human adu
29	141	4.1	184	2	AAH86031	Human sin
30	140	4.1	243	4	AAS57938	CDNA #614
31	135	4.0	1707	7	ACA90176	CDNA enco
32	129	3.8	228	3	AAA41968	Human sec
33	93	2.7	253	4	AAS58053	CDNA #729
34	92	2.7	598	5	ABV56560	Human pro
35	83	2.4	380	6	ABL67315	Thyroid c
36	83	2.4	380	6	ABN94637	Gene #113
37	68	2.0	412	6	ABL64217	Stomach c
38	62	1.8	145	4	AAS26568	Human CDN
39	62	1.8	145	7	ABX73909	Human nov
40	56	1.6	1078	3	AAA02581	Human col
41	55	1.6	564	8	AAI57525	Human IGF
42	42	1.2	419	4	AAI86673	Human pol
43	37	1.1	495	8	ACH39118	Human foe
44	27	0.8	1404	6	ABQ22104	Oligonuc1
45	27	0.8	1404	6	ABQ22105	Oligonuc1

ALIGNMENTS

RESULT 1
AAZ36152
ID AAZ36152 standard; DNA; 3412 BP.
XX
AC AAZ36152;
XX
DT 11-FEB-2000 (first entry)
XX
DE DNA encoding cancer associated antigen KOC-3.
XX
KW Cancer associated antigen; KOC-3; cancer; vaccine; CT7; ss.
XX
OS Homo sapiens.
XX
PN WO9954738-A1.
XX
PD 28-OCT-1999.
XX
PF 16-MAR-1999; 99WO-US005766.
XX
PR 17-APR-1998; 98US-00061709.
XX
PA (LUDW-) LUDWIG INST CANCER RES.

Chen Y, Gure A, Tsang S, Stockert E, Jager E, Knuth A, Old LJ;
WPI; 2000-013284/01.
Nucleotides representing cancer-associated genes, used to develop products for the diagnosis, monitoring and treatment of cancers.
Claim 55; Page 41; 44pp; English.
The present sequence represents a cancer associated antigen gene designated KOC-3. The specification also describes a cancer associated antigen designated CT7. The CT7 polynucleotide was isolated from SK-MEL-37 melanoma cells. The polypeptide has some homology with MAGS-10, a limited to about 210 carboxy terminal amino acids. The amino terminal of the protein has a repetitive pattern, with repeats rich in serine, proline, glutamine and leucine, and an almost invariable core of the peptide given in AA43877. The CT7 polypeptide can be processed to peptides which provoke lysis by cytolytic T cells. The polynucleotides and polypeptides can be used for treating a cancerous condition and screening for or diagnosing cancerous conditions. The cancer associated antigens can be used as an immunogenic or vaccine composition with an adjuvant, e.g. a cytokine, a saponin, or granulocyte macrophage-colony

PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226689P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 12-SEP-2000; 2000US-0232081P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.

PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 05-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-488783/53.
XX P-PSDB; AAU16163.
XX
PT New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives.
XX
FS Claim 1; SEQ ID NO 329; 980pp; English.
XX
CC The invention relates to isolated nucleic acid molecules and their
CC encoded secreted proteins. The nucleic acids and proteins are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. Antibodies to the proteins can also be used in
CC alleviating symptoms associated with the disorders and in diagnostic
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays
CC (ELISA). Disorders which are diagnosed or treated include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
CC and many other disorders listed in the specification. The polypeptides
CC can also be used to aid wound healing and epithelial cell proliferation,
CC to prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence encodes a novel secreted protein of the invention. Note: The

CC sequence data for this patent did not form part of the printed									
Query Match 86.5%; Score 2953; DB 4; Length 3694;									
Best Local Similarity 99.9%; Pred. No. 0;									
Matches 3303; Conservative 0; Mismatches 1; Indels 3; Gaps 2;									
Qy	68	GGATGATGACAAAGCTTTACATCGGGAACCTGAGACCCCGCGTCACCGCGACGACCTCC	127						
Db	91	GGATGATGACAAAGCTTTACATCGGGAACCTGAGACCCCGCGTCACCGCGACGACCTCC	150						
Qy	128	GGCAGCTCTTTGGGGACAGAAAGCTGCCCTTGGCGGACAGGTCCTCTCAAGTCCGGCT	187						
Db	151	GGCAGCTCTTTGGGGACAGAAAGCTGCCCTTGGCGGACAGGTCCTCTCAAGTCCGGCT	210						
Qy	188	ACGCCCTTGTGGACTACCCCGACAGAACTGGGCCATCCCGGCCATCGACACCTCTCGG	247						
Db	211	ACGCCCTTGTGGACTACCCCGACAGAACTGGGCCATCCCGGCCATCGACACCTCTCGG	270						
Qy	248	GTAAAGTGGAAATTCATGCGGAATCATGGAAGTTGATTAATCAGTCTCTTAAAGCTAA	307						
Db	271	GTAAAGTGGAAATTCATGCGGAATCATGGAAGTTGATTAATCAGTCTCTTAAAGCTAA	330						
Qy	308	GGAGCAGGAAATTCAGATTGGAATCGAAACATCCCTCCTCACCTGCGAGTGGAGGTGTGGATG	367						
Db	331	GGAGCAGGAAATTCAGATTGGAATCGAAACATCCCTCCTCACCTGCGAGTGGAGGTGTGGATG	390						
Qy	368	GACTTTTGGTCAATATGGGACAGTGGGAATGGAACAAAGTCAACACGACACAGAAA	427						
Db	391	GACTTTTGGTCAATATGGGACAGTGGGAATGGAACAAAGTCAACACGACACAGAAA	450						
Qy	428	CCGCCGTGTCACAGTCACATATGCAACAGAGNAGAGCAAAATAGCCATGGAGAGC	487						
Db	451	CCGCCGTGTCACAGTCACATATGCAACAGAGNAGAGCAAAATAGCCATGGAGAGC	510						
Qy	488	TAAGCGGGCATCAGTTTGAAACTACCTCCCTCAAGATTTCTACATCCCGATGAAGG	547						
Db	511	TAAGCGGGCATCAGTTTGAAACTACCTCCCTCAAGATTTCTACATCCCGATGAAGG	570						
Qy	548	TGAGTCCCTTGCCTCCCTCAGGAGCCAGGTCGGGACCACTCTTCCCGGAGCAAG	607						
Db	571	TGAGTCCCTTGCCTCCCTCAGGAGCCAGGTCGGGACCACTCTTCCCGGAGCAAG	630						
Qy	608	GCCACGCCCTTGGGGGCACTTCTCAGGCGACAGATTGATTTCCCGCTCGGATCCTGG	667						
Db	631	GCCACGCCCTTGGGGGCACTTCTCAGGCGACAGATTGATTTCCCGCTCGGATCCTGG	690						
Qy	668	TCCCCACCCAGTTTGTGGTGCCATCATCGAAAGAGGGCTTGACCATAAAGAACATCA	727						
Db	691	TCCCCACCCAGTTTGTGGTGCCATCATCGAAAGAGGGCTTGACCATAAAGAACATCA	750						
Qy	728	CTAAGCAGACCCAGTCCCGGGTAGATATCCATGAAAGAGAACTCTGGAGCTGCAGAGA	787						
Db	751	CTAAGCAGACCCAGTCCCGGGTAGATATCCATGAAAGAGAACTCTGGAGCTGCAGAGA	810						
Qy	788	AGCTGTCCACCTCCATGCGCCAGGCGGACTTCTGAGCATGCGCATGATCTTCTG	847						
Db	811	AGCTGTCCACCTCCATGCGCCAGGCGGACTTCTGAGCATGCGCATGATCTTCTG	870						
Qy	848	AAATCATGCAAAAGAGGAGATGAGTGGGAGGAGATTCCTCTGAAATCT	907						
Db	871	AAATCATGCAAAAGAGGAGATGAGTGGGAGGAGATTCCTCTGAAATCT	930						
Qy	908	TGGCACAATATGGCTTGGGAAGATGATGGAAGAGAGGAGAGAAATTTGAAGAAA	967						
Db	931	TGGCACAATATGGCTTGGGAAGATGATGGAAGAGAGGAGAGAAATTTGAAGAAA	990						
Qy	968	TTGAACTGAAACAGGAGCAGATGAGTGGGAGGAGTTCCTCTGAAATCT	1027						
Db	991	TTGAACTGAAACAGGAGCAGATGAGTGGGAGGAGTTCCTCTGAAATCT	1050						
Qy	1028	ACCCGGAAGAACCATCAGTGTGAAGGCGACAGTTGAGGCTGTGCCAGTCTGAGATAG	1087						
Db	1051	ACCCGGAAGAACCATCAGTGTGAAGGCGACAGTTGAGGCTGTGCCAGTCTGAGATAG	1110						

Qy	1088	AGATTATGAAGAGCTGCGTGAGGCTTTGAAATGATATGCTGGCTGTTAAACAAG	1147
Db	1111	AGATTATGAAGAGCTGCGTGAGGCTTTGAAATGATATGCTGGCTGTTAAACAAG	1170
Qy	1148	CCATCTGATCCAGAGGTTGAACCTCAGCGCACTTGGCATCTTTTCAACAGGAGTGTCCG	1207
Db	1171	CCATCTGATCCAGAGGTTGAACCTCAGCGCACTTGGCATCTTTTCAACAGGAGTGTCCG	1230
Qy	1208	TGCTATCTCCACAGAGGCCCCCGGGAGCTCCCCCGCTGCCCTTACACCCCTTCA	1267
Db	1231	TGCTATCTCCACAGAGGCCCCCGGGAGCTCCCCCGCTGCCCTTACACCCCTTCA	1290
Qy	1268	CTACCCACTCCGATATCTTCTCAGGCTGTACCCCATCAACAGTTTGGCCGCTTCCCG	1327
Db	1291	CTACCCACTCCGATATCTTCTCAGGCTGTACCCCATCAACAGTTTGGCCGCTTCCCG	1350
Qy	1328	ATCATCACTCTTATCCAGAGCAGGAGATTGTGAATCTCTTCATCCCAACCCAGGCTGTGG	1387
Db	1351	ATCATCACTCTTATCCAGAGCAGGAGATTGTGAATCTCTTCATCCCAACCCAGGCTGTGG	1410
Qy	1388	GCGCCATCATCGGGAAGAGGGGGCCACACATCAAAAGAGCTGGCGAGATTCCCGGAGCCT	1447
Db	1411	GCGCCATCATCGGGAAGAGGGGGCCACACATCAAAAGAGCTGGCGAGATTCCCGGAGCCT	1470
Qy	1448	CTATCAAGATTGCCCCCTGCGGAAGGCCAGACGTCAAGCGAAAGAGTGGTCAATCACC	1507
Db	1471	CTATCAAGATTGCCCCCTGCGGAAGGCCAGACGTCAAGCGAAAGAGTGGTCAATCACC	1530
Qy	1508	GGCCACCGGAAGCCAGTTCAGAGCCAGGAGCGGATCTTTGGGAACCTGAAAGAGAAA	1567
Db	1531	GGCCACCGGAAGCCAGTTCAGAGCCAGGAGCGGATCTTTGGGAACCTGAAAGAGAAA	1590
Qy	1568	ACTTCTTTTAAACCCCAAGAGAGTGAAGCTGGAAGCGCATATCAGAGTGCCTCTTCCA	1627
Db	1591	ACTTCTTTTAAACCCCAAGAGAGTGAAGCTGGAAGCGCATATCAGAGTGCCTCTTCCA	1650
Qy	1628	CAGTGGCCGGGTGATTGGCAAGAGTGGCAAGACCGGTGAACGAACTGAACCA	1687
Db	1651	CAGTGGCCGGGTGATTGGCAAGAGTGGCAAGACCGGTGAACGAACTGAACCA	1710
Qy	1688	GTCCAGAGTCATCTGCTCGCTGACCAAAAGCGAGTGAATGAAGTGAAGTGAATCTCA	1747
Db	1711	GTCCAGAGTCATCTGCTCGCTGACCAAAAGCGAGTGAATGAAGTGAAGTGAATCTCA	1770
Qy	1748	GAATTATCGGGCACTTCTTGTAGCCAGACTGCACAGCGCAAGATCAGGAAATTTCTAC	1807
Db	1771	GAATTATCGGGCACTTCTTGTAGCCAGACTGCACAGCGCAAGATCAGGAAATTTCTAC	1830
Qy	1808	AAAGTGTGAAGCAGGAGCAGAAATACCTCAGGAGTGCCTTCAAGCGCAGCAAGT	1867
Db	1831	AAAGTGTGAAGCAGGAGCAGAAATACCTCAGGAGTGCCTTCAAGCGCAGCAAGT	1890
Qy	1868	GAGCTCCACAGCAGCCAGCAAAACAAAGATGAATGTAGCCCTTCCCAACCTGACAG	1927
Db	1891	GAGCTCCACAGCAGCCAGCAAAACAAAGATGAATGTAGCCCTTCCCAACCTGACAG	1950
Qy	1928	AATGAGACCAAAACCGAGCCAGATCGGAGCAAAACCAAGACCATCTGAGGAATGAG	1987
Db	1951	AATGAGACCAAAACCGAGCCAGATCGGAGCAAAACCAAGACCATCTGAGGAATGAG	2010
Qy	1988	AAGTCTCGGAGCGGCGAGGAGCTCTGCGAGGCTGTGAGAACCCAGGGGCGGAGGAG	2047
Db	2011	AAGTCTCGGAGCGGCGAGGAGCTCTGCGAGGCTGTGAGAACCCAGGGGCGGAGGAG	2070
Qy	2048	GGGCGGGAGGTCAGCCAGTTTGGCAGAACCAACCGAGCCCGGCTTCCCGCCCCCAGG	2107
Db	2071	GGGCGGGAGGTCAGCCAGTTTGGCAGAACCAACCGAGCCCGGCTTCCCGCCCCCAGG	2130
Qy	2108	GCTTCTGAGGCTTCAAGCCATCCACTTCACTTCACTCCGATCTCTCTGAACTCCAC	2167
Db	2131	GCTTCTGAGGCTTCAAGCCATCCACTTCACTTCACTCCGATCTCTCTGAACTCCAC	2190

QY	2168	GACGCTATCCCTTTTGTAGTGAACATACTAGTGAAACGTTGTCAAGGCCAAGCAAAATGC	2227	Db	3271	AGCATTTTTTTTATTTTTTAAATAAAATGAGTTGGAAAAAATAAGATAATCAACTGCCA	3330
Db	2191	GACGCTATCCCTTTTGTAGTGAACATACTAGTGAAACGTTGTCAAGGCCAAGCAAAATGC	2250	QY	3305	GCCTGGAGAAAGTGACAGTGCACAGTGTGCAACAGCTGTTCTGAATTGTCTTCCGCTAGCC	3364
QY	2228	ACACCCCTTTTCTGTGGCAAAATCGTCTCTGTACATGTGTGTACATATTAGAAAGGGAAGA	2287	Db	3331	GCCTGGAGAAAGTGACAGTGCACAGTGCACAGCTGTTCTGAATTGTCTTCCGCTAGCC	3390
Db	2251	ACACCCCTTTTCTGTGGCAAAATCGTCTCTGTACATGTGTGTACATATTAGAAAGGGAAGA	2310	QY	3365	AAGAAC 3371	
QY	2288	TGTTAAGATATGTGGCTGTGGTTTACACAGGGTGCCTGCAGCGTAAATATATTATTAGAA	2347	Db	3391	AAGAAC 3397	
Db	2311	TGTTAAGATATGTGGCTGTGGTTTACACAGGGTGCCTGCAGCGTAAATATATTATTAGAA	2370	RESULT 3			
QY	2348	ATAATATATCAATAACTCAACTCAACTCAATTTTTTAATCAATTAATTTTTTTTCT	2407	ABX73491			
Db	2371	ATAATATATCAATAACTCAACTCAACTCAATTTTTTAATCAATTAATTTTTTTTCT	2430	ID ABX73491 standard; DNA; 3694 BP.			
QY	2408	TTTTAAAGAGAAAGCAGGCTTTTCTAGACTTTTAAAGAAATTAAGTCTTTGGAGGCTCAC	2467	XX ABX73491;			
Db	2431	TTTTAAAGAGAAAGCAGGCTTTTCTAGACTTTTAAAGAAATTAAGTCTTTGGAGGCTCAC	2490	XX			
QY	2468	GGTGTAGAGAGAGCTTTGAGGCCACCGGCACAAAATTCACCCAGAGGGAATCTCGTCG	2527	DT 18-MAR-2003 (first entry)			
Db	2491	GGTGTAGAGAGAGCTTTGAGGCCACCGGCACAAAATTCACCCAGAGGGAATCTCGTCG	2550	XX Human novel polynucleotide #319.			
QY	2528	GAAGACACTCACGCGAGTCTGGATCACTGTGTATGTCAACAGAAGGGATACCGTCTC	2587	DE			
Db	2551	GAAGACACTCACGCGAGTCTGGATCACTGTGTATGTCAACAGAAGGGATACCGTCTC	2610	XX Human; gene; ds; neural disorder; immune system disorder; renal disorder;			
QY	2588	CTTGAAGAGGAACCTGTCACTCCTCATGCTGTCTAGCTCATACACCACTTCTCTTT	2647	KW muscular disorder; respiratory disease; reproductive disorder;			
Db	2611	CTTGAAGAGGAACCTGTCACTCCTCATGCTGTCTAGCTCATACACCACTTCTCTTT	2670	KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;			
QY	2648	GCTTCACAGGTTTTAAACTGGTTTTTTCATCTCTATATAATTCCTGTCTCTCTCTG	2707	KW hyperproliferative disorder; inflammatory disease; allergic reaction;			
Db	2671	GCTTCACAGGTTTTAAACTGGTTTTTTCATCTCTATATAATTCCTGTCTCTCTCTG	2730	KW blood related disorder; cancer; immunosuppressive; antiinflammatory;			
QY	2708	TTTATCTCTCCCTCCCTCCCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2767	KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;			
Db	2731	TTTATCTCTCCCTCCCTCCCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2790	XX haemostatic; antiarteriosclerotic.			
QY	2768	ATCCCTCCATCTCAATCCGATCTACGCA--CCCCCCCCCCCCCAGGCAAGCATGCT	2825	OS Homo sapiens.			
Db	2791	ATCCCTCCATCTCAATCCGATCTACGCA--CCCCCCCCCCCCCAGGCAAGCATGCT	2850	PN US2002132753-A1.			
QY	2826	CTGAGTATCATCATCACAAAGGAACAAAGCGAAACACAAACAGCGCTCAACTTAC	2885	XX 19-SEP-2002.			
Db	2851	CTGAGTATCATCATCACAAAGGAACAAAGCGAAACACAAACAGCGCTCAACTTAC	2910	PF 17-JAN-2001; 2001US-00764864.			
QY	2886	ACTTGGTTACTCAAAAGAACAAAGTCAATGGTACTTCTCTCAATCTCAATCTTGAATTCCTC	2945	XX 31-JAN-2000; 2000US-0179065P.			
Db	2911	ACTTGGTTACTCAAAAGAACAAAGTCAATGGTACTTCTCTCAATCTCAATCTTGAATTCCTC	2970	PR 04-FEB-2000; 2000US-0180628P.			
QY	2946	AACAGGAACCCACCAACCAACCAATCAACCAAGAAATAATCCCAATCAAGA	3005	PR 28-JUN-2000; 2000US-0214886P.			
Db	2971	AACAGGAACCCACCAACCAACCAATCAACCAAGAAATAATCCCAATCAAGA	3030	PR 07-JUL-2000; 2000US-0216647P.			
QY	3006	ATGTAATTTGTCTTTTGCATTTTGGTGTATAAGCCATCAATATTCAGCAAAATGATTC	3065	PR 07-JUL-2000; 2000US-0216880P.			
Db	3031	ATGTAATTTGTCTTTTGCATTTTGGTGTATAAGCCATCAATATTCAGCAAAATGATTC	3090	PR 11-JUL-2000; 2000US-0217487P.			
QY	3066	TTTCTTTT-AAAAAAAATGTGGAGAAAGTAGAAATTTACCAAGTTTGTGGCCGAG	3124	PR 14-JUL-2000; 2000US-0218290P.			
Db	3091	TTTCTTTAAAAAAAATGTGGAGAAAGTAGAAATTTACCAAGTTTGTGGCCGAG	3150	PR 26-JUL-2000; 2000US-0220963P.			
QY	3125	CGGTTAAATTCACAGATTTTTTAAAGAAATAACACAGAAAGTACTCAGGTGT	3184	PR 14-AUG-2000; 2000US-0225447P.			
Db	3151	CGGTTAAATTCACAGATTTTTTAAAGAAATAACACAGAAAGTACTCAGGTGT	3210	PR 14-AUG-2000; 2000US-0225757P.			
QY	3185	TTTACCTCAGCACCTTCTCTGTGTTCCTTAGAGATTTTGTAAAGCTGATAGTTGG	3244	PR 22-AUG-2000; 2000US-0225758P.			
Db	3211	TTTACCTCAGCACCTTCTCTGTGTTCCTTAGAGATTTTGTAAAGCTGATAGTTGG	3270	PR 30-AUG-2000; 2000US-0228924P.			
QY	3245	AGCATTTTTTTTATTTTAAATAAAATGAGTTGGAAAAAATAAGATATCAACTGCCA	3304	PR 01-SEP-2000; 2000US-0229287P.			
				PR 01-SEP-2000; 2000US-0229343P.			
				PR 01-SEP-2000; 2000US-0229345P.			
				PR 05-SEP-2000; 2000US-0229509P.			
				PR 05-SEP-2000; 2000US-0229513P.			
				PR 08-SEP-2000; 2000US-0231413P.			
				PR 21-SEP-2000; 2000US-0234223P.			
				PR 21-SEP-2000; 2000US-0234274P.			
				PR 25-SEP-2000; 2000US-0234997P.			
				PR 27-SEP-2000; 2000US-0235834P.			
				PR 29-SEP-2000; 2000US-0236327P.			
				PR 29-SEP-2000; 2000US-0236367P.			
				PR 29-SEP-2000; 2000US-0236368P.			
				PR 29-SEP-2000; 2000US-0236369P.			

PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX
(ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
PI Rosen CA, Ruben SM, Barash SC;
DR
DR WPI; 2003-147444/14.
DR P-PSDB; ABUS5231.
XX
XX
PT New polypeptides and nucleic acids, useful in gene therapy for treating,
PT inhibiting or preventing e.g. neural, immune system, muscular,
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
PT renal disorders.
XX
PS Claim 1; SEQ ID NO 329; 402pp; English.
XX
CC The invention relates to human novel polypeptides and their associated
CC polynucleotides. The polypeptides and polynucleotides are useful in gene
CC therapy for treating, inhibiting or preventing neural disorders, immune
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
CC leukemia), inflammatory diseases (e.g. septic shock, bursitis and
CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
CC infarction) and cancerous diseases. Sequences ABX73173-ABX74167 represent
CC human novel polynucleotides of the invention
XX
SQ Sequence 3694 BP; 1103 A; 934 C; 839 G; 817 T; 0 U; 1 Other;

Query Match 86.5%; Score 2953; DB 7; Length 3694;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3303; Conservative 0; Mismatches 1; Indels 3; Gaps 2;

QY 68 GGATGATGAACAGCTTTACATCGGGAACTCTGAGCCCGCGGTACCGCCGACAGCTCC 127
DB 1111 AGATTATGAAGAAAGCTCGGTGAGGCGCTTTGAAAATGATATGCTGGCTGTAAACCAAG 1170
QY 128 GGCAGCTCTTTGGGGACAGAAAGCTGCCCTTGGCGGACAGGTCTCTGCTGAAGTCCGGCT 187
DB 151 GGCAGCTCTTTGGGGACAGAAAGCTGCCCTTGGCGGACAGGTCTCTGCTGAAGTCCGGCT 210
QY 188 ACGCTTCTGCGGACTACCGGACAGAACTGGCGGCATCCGGGCCATCGAGACCTCTTCGG 247
DB 211 ACGCTTCTGCGGACTACCGGACAGAACTGGCGGCATCCGGGCCATCGAGACCTCTTCGG 270
QY 248 GTAAGTGGAAATTCATGGGAATCATGGAGTTGATCTCAGTCTCTAAAAGCTAA 307
DB 271 GTAAGTGGAAATTCATGGGAATCATGGAGTTGATCTCAGTCTCTAAAAGCTAA 330
QY 308 GGAGCAGGAAATTCAGATTCGAACATCCCTCTCAGCTGCGAGTGGAGGTGTGGATG 367
DB 331 GGAGCAGGAAATTCAGATTCGAACATCCCTCTCAGCTGCGAGTGGAGGTGTGGATG 390

QY 368 GACTTTTGGCTCAATATGGGACAGTGGAGAAATGTGAAACAAGTCAACACAGACACAGAAA 427
DB 391 GACTTTTGGCTCAATATGGGACAGTGGAGAAATGTGAAACAAGTCAACACAGACACAGAAA 450
QY 428 CGCGCGTTGTACAGTTCACATATGCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 487
DB 451 CGCGCGTTGTCAACAGTTCACATATGCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 510
QY 488 TAAGCGGGCATCAGTTTGGAGAACTACTCTCTTCAAGATTTTCTTACATCCCGGATGAAGAG 547
DB 511 TAAGCGGGCATCAGTTTGGAGAACTACTCTCTTCAAGATTTTCTTACATCCCGGATGAAGAG 570
QY 548 TGAGTCCCTTTCGCCCCCTTCAGCGAGCCCGTGGGAGACCACTCTTTCGCGGAGCAAG 607
DB 571 TGAGTCCCTTTCGCCCCCTTCAGCGAGCCCGTGGGAGACCACTCTTTCGCGGAGCAAG 630
QY 608 GCCACGCCCTGGGGGACCTTCTCAGGCCAGACAGATTTGATTTCCCGCTGCGGATCCTGG 667
DB 631 GCCACGCCCTGGGGGACCTTCTCAGGCCAGACAGATTTGATTTCCCGCTGCGGATCCTGG 690
QY 668 TCCCCACCCAGTTTGTGGTGGCCATCATCGGAAAGGAGGCTTGAACATAAAGAAATCA 727
DB 691 TCCCCACCCAGTTTGTGGTGGCCATCATCGGAAAGGAGGCTTGAACATAAAGAAATCA 750
QY 728 CTAAGCAGACCCAGTCCCGGTTAGATATCCATAGAAAAGAGAACTCTGGAGCTGCGAGAG 787
DB 751 CTAAGCAGACCCAGTCCCGGTTAGATATCCATAGAAAAGAGAACTCTGGAGCTGCGAGAG 810
QY 788 AGCTGTCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCC 847
DB 811 AGCTGTCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCC 870
QY 848 AAATCATGAGAAAGAGGAGATGAGACCAAACTAGCCGAGAGATTCCTCTGAAAATCT 907
DB 871 AAATCATGAGAAAGAGGAGATGAGACCAAACTAGCCGAGAGATTCCTCTGAAAATCT 930
QY 908 TGCCACAAATGCTTGGTGGAGAGTATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 967
DB 931 TGCCACAAATGCTTGGTGGAGAGTATGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 990
QY 968 TTCAATGAAACAGGAGGACCAAGATACATCTCTTTTGGAGGATTTTGGAGATATACA 1027
DB 991 TTGAAATGAAACAGGAGGACCAAGATACATCTCTTTTGGAGGATTTTGGAGATATACA 1050
QY 1028 ACCCGAAAGAAACCATCTGCTGAAAGGACAGCTTGAGGCGCTGTCGAGTGTGAGATAG 1087
DB 1051 ACCCGAAAGAAACCATCTGCTGAAAGGACAGCTTGAGGCGCTGTCGAGTGTGAGATAG 1110
QY 1088 AGATTATGAAGAAAGCTCGGTGAGGCGCTTTGAAAATGATATGCTGGCTGTAAACCAAG 1147
DB 1111 AGATTATGAAGAAAGCTCGGTGAGGCGCTTTGAAAATGATATGCTGGCTGTAAACCAAG 1170
QY 1148 CCAATCTGATCCAGAGGTGAACTCAGGSCACTTGGSCATCTTTTCAACAGAGACTGTCCG 1207
DB 1171 CCAATCTGATCCAGAGGTGAACTCAGGSCACTTGGSCATCTTTTCAACAGAGACTGTCCG 1230
QY 1208 TGCTATCTCCACAGAGGCGCCCGCGGAGCTCCCCCGCTGCCCTTACACCCCTTCA 1267
DB 1231 TGCTATCTCCACAGAGGCGCCCGCGGAGCTCCCCCGCTGCCCTTACACCCCTTCA 1290
QY 1268 CTAACCACTCCGAGTACTTCTCAGGCTGTACCCCATACAGTTTGGCGCTTCCCGC 1327
DB 1291 CTAACCACTCCGAGTACTTCTCAGGCTGTACCCCATACAGTTTGGCGCTTCCCGC 1350
QY 1328 ATCATCACTCTTATCCAGAGCAGAGATTTGTGAATCTCTTCAACCAAGAGAGAGAGAG 1387
DB 1351 ATCATCACTCTTATCCAGAGCAGAGATTTGTGAATCTCTTCAACCAAGAGAGAGAGAG 1410
QY 1388 GCGCCATCATCGGAAAG 1447
DB 1411 GCGCCATCATCGGAAAG 1470

QY	1448	CTATCAAGATTGCCCTCGGAAAGCCAGACGTACAGGAAGAAGTGGTCATCATCACCG	1507
Db	1471	CTATCAAGATTGCCCTCGGAAAGCCAGACGTACAGGAAGAAGTGGTCATCATCACCG	1530
QY	1508	GGCCACCGAAGCCAGTTCAAGGCCCAGGAGCGGATCTTTGGGAAAACCTGAAAGAGGAAA	1567
Db	1531	GGCCACCGAAGCCAGTTCAAGGCCCAGGAGCGGATCTTTGGGAAAACCTGAAAGAGGAAA	1590
QY	1568	ACTTCTTTAACCCCAAAGAAAGTAGTGGAAGCTGGAAGCGCATATCAGAGTGCCTCTTCCA	1627
Db	1591	ACTTCTTTTAACCCCAAAGAAAGTAGTGGAAGCTGGAAGCGCATATCAGAGTGCCTCTTCCA	1650
QY	1628	CAGCTGGCGGGTGAATTGGCAAAAGGTGGCAAGACCGTGAACGACTGACAGACTTAAACCA	1687
Db	1651	CAGCTGGCGGGTGAATTGGCAAAAGGTGGCAAGACCGTGAACGACTGACAGACTTAAACCA	1710
QY	1688	GTGCAGAAGTCATCGTGCTCGTGACCACAAACGCCCAGATGAAAAATGAGGAAAGTGATCGTCA	1747
Db	1711	GTGCAGAAGTCATCGTGCTCGTGACCACAAACGCCCAGATGAAAAATGAGGAAAGTGATCGTCA	1770
QY	1748	GAAATTATCGGSCACTCTTTGCTAGCCAGACTGACAGCGCAAGATCAGGGAATTTGTAC	1807
Db	1771	GAAATTATCGGSCACTCTTTGCTAGCCAGACTGACAGCGCAAGATCAGGGAATTTGTAC	1830
QY	1808	AACAGGTGAAGCAGCAGCAGAGCAGAAAAATAACCTCAGGGAGTGCCTCACGCGCAGCAAGT	1867
Db	1831	AACAGGTGAAGCAGCAGCAGAGCAGAAAAATAACCTCAGGGAGTGCCTCACGCGCAGCAAGT	1890
QY	1868	GAGCTCCCA CAGGCACA GAGCAAAACAAACGATGAATGTAGCCCTTCCAA CACTGACAG	1927
Db	1891	GAGCTCCCA CAGGCACA GAGCAAAACAAACGATGAATGTAGCCCTTCCAA CACTGACAG	1950
QY	1928	AATGAGACCAAACGAGCCAGCAGATCGGAGCAAAACCAACCAAGACCAATCTGAGGAATGAG	1987
Db	1951	AATGAGACCAAACGAGCCAGCAGATCGGAGCAAAACCAACCAAGACCAATCTGAGGAATGAG	2010
QY	1988	AAGTCTCGGAGGCGGCAGGGA CTCTGCCAGGCCCTGAGAA CCCCAGGGCGCAGGAG	2047
Db	2011	AAGTCTCGGAGGCGGCAGGGA CTCTGCCAGGCCCTGAGAA CCCCAGGGCGCAGGAG	2070
QY	2048	GGGCGGGGAAGGTCAGCCAGGTTTTGCAGAACCAACGAGCCCGCCCTCCCGCCCCCAGG	2107
Db	2071	GGGCGGGGAAGTCAGCCAGGTTTTGCAGAACCAACGAGCCCGCCCTCCCGCCCCCAGG	2130
QY	2108	GCTTCTCAGGCTTCAGCCATCCACTTCAACATCCACTCGGATCTCTCTGAACTCCAC	2167
Db	2131	GCTTCTCAGGCTTCAGCCATCCACTTCAACATCCACTCGGATCTCTCTGAACTCCAC	2190
QY	2168	GAGCTATCCCTTTAGTTGNACTAA CATAGGTGAACGTGTTCAAGCCCAAGCAAAATGC	2227
Db	2191	GAGCTATCCCTTTAGTTGNACTAA CATAGGTGAACGTGTTCAAGCCCAAGCAAAATGC	2250
QY	2228	ACACCTTTTCTGTGGCAATCGTCTCTGTACATGTGTACATATTAGAAAGGGAAGA	2287
Db	2251	ACACCTTTTCTGTGGCAATCGTCTCTGTACATGTGTACATATTAGAAAGGGAAGA	2310
QY	2288	TGTTAAGATATGGCCCTGTGGTTACAGAGGTGCCTGACGGTAAATATATTTAGAA	2347
Db	2311	TGTTAAGATATGGCCCTGTGGTTACAGAGGTGCCTGACGGTAAATATATTTAGAA	2370
QY	2348	ATAATATCAAAATAACTCAACTAACCTCCAAATTTTAAATCAATATTTTCTTCT	2407
Db	2371	ATAATATCAAAATAACTCAACTAACCTCCAAATTTTAAATCAATATTTTCTTCT	2430
QY	2408	TTTTAAAGAAAGCAGGCTTTCTAGACTTTTAAAGATAAAGTCTTTGGAGGCTCTCAC	2467
Db	2431	TTTTAAAGAAAGCAGGCTTTCTAGACTTTTAAAGATAAAGTCTTTGGAGGCTCTCAC	2490
QY	2468	GGTGTAGAGGAGCTTTTGGGCCACCCGACAAAAATTCACCCAGAGGGAATCTCGTGC	2527
Db	2491	GGTGTAGAGGAGCTTTTGGGCCACCCGACAAAAATTCACCCAGAGGGAATCTCGTGC	2550
QY	2528	GAAGGACACTCAGGCAAGTCTGGATCACTGTGTATGTCAACAGAGGAGATACCGTCTC	2587

XX WO954738-A1.
XX 28-OCT-1999.
XX 16-MAR-1999; 99WO-US005766.
XX 17-APR-1998; 98US-00061709.
XX (LUDW-) LUDWIG INST CANCER RES.
XX Chen Y, Gure A, Tsang S, Stockert E, Jager E, Knuth A, Old LJ;
XX WPI; 2000-013284/01.
XX
XX Nucleotides representing cancer-associated genes, used to develop
XX products for the diagnosis, monitoring and treatment of cancers.
XX
XX Claim 55; Page 42-43; 44pp; English.
XX
XX The present sequence represents an alternative form of a cancer
XX associated antigen gene designated KOC-3. The specification also
XX describes a cancer associated antigen designated CT7. The CT7
XX polynucleotide was isolated from SK-MEL-37 melanoma cells. The
XX polypeptide has some homology with MAGE-10, limited to about 210 carboxy
XX terminal amino acids. The amino terminal of the protein has a repetitive
XX pattern, with repeats rich in serine, proline, glutamine and leucine, and
XX an almost invariable core of the peptide given in AAY43877. The CT7
XX polypeptide can be processed to peptides which provoke lysis by cytolytic
XX T cells. The polynucleotides and polypeptides can be used for treating a
XX cancerous condition and screening for or diagnosing cancerous conditions.
XX The cancer associated antigens can be used as an immunogenic or vaccine
XX composition with an adjuvant, e.g. a cytokine, a saponin, or granulocyte
XX macrophage-colony stimulating factor (GM-CSF)
XX
XX Sequence 3283 BP; 945 A; 833 C; 779 G; 725 T; 0 U; 1 Other;
XX
XX Query Match 62.8%; Score 2143; DB 3; Length 3283;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 2143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1270 ACCCACTCCGATATCTTCAGGCTGTACCCCATCACAGTTTGGCCGTTCCCGCAT 1329
DB 1141 ACCCACTCCGATATCTTCAGGCTGTACCCCATCACAGTTTGGCCGTTCCCGCAT 1200
QY 1330 CATCACTCTTATCCAGAGCAGGAGATTGTGAATCTCTTATCCCAACCCAGGCTGTGGGC 1389
DB 1201 CATCACTCTTATCCAGAGCAGGAGATTGTGAATCTCTTATCCCAACCCAGGCTGTGGGC 1260
QY 1390 GCATCATCCGGAGAGAGGGGACACATCAACAGCTGGGAGATTGCGCGAGCCTCT 1449
DB 1261 GCCATCATCCGGAGAGAGGGGACACATCAACAGCTGGGAGATTGCGCGAGCCTCT 1320
QY 1450 ATCAAGATTGCCCTCCGGAAGGCCAGAGCTCAGCGAAGGATGCTCATCATCACCGGG 1509
DB 1321 ATCAAGATTGCCCTCCGGAAGGCCAGAGCTCAGCGAAGGATGCTCATCATCACCGGG 1380
QY 1510 CCACCGAAGCCCAAGTTCAAGGCCCAAGGACGATCTTTGGGAACTGAAAGAGGAAAC 1569
DB 1381 CCACCGAAGCCCAAGTTCAAGGCCCAAGGACGATCTTTGGGAACTGAAAGAGGAAAC 1440
QY 1570 TTCTTTAACCCCAAGAGAGAGTGAAGTGAAGCGCATATCAGAGTGCCTCTTCCACA 1629
DB 1441 TTCTTTAACCCCAAGAGAGAGTGAAGTGAAGCGCATATCAGAGTGCCTCTTCCACA 1500
QY 1630 GCTGGCCGGGTGATTGGCAAAGGTGGCAAGACCGTGAACGAACTGCAAGCTTAACAGT 1689
DB 1501 GCTGGCCGGGTGATTGGCAAAGGTGGCAAGACCGTGAACGAACTGCAAGCTTAACAGT 1560
QY 1690 GCAGAAGTCACTGTCGCTCGTGAACCAACGCCAGATGAAATGAGGAGTGAATCGTCAGA 1749
DB 1561 GCAGAAGTCACTGTCGCTCGTGAACCAACGCCAGATGAAATGAGGAGTGAATCGTCAGA 1620

QY 1750 ATTATCGGCGACTTCTTTGCTAGCAGACTGACAGCGAAGATCAGGAAATTTGTACAA 1809
DB 1621 ATTATCGGCGACTTCTTTGCTAGCAGACTGACAGCGAAGATCAGGAAATTTGTACAA 1680
QY 1810 CAGGTGAAGCAGCAGGAGCAGAAATACCTCTCAGGAGTCCCTCTCAGACGCGCAAGTGA 1869
DB 1681 CAGGTGAAGCAGCAGGAGCAGAAATACCTCTCAGGAGTCCCTCTCAGACGCGCAAGTGA 1740
QY 1870 GGCTCCACAGCAGCAGCAGCAAAACAAACGAGTGAATGTAGCCCTTCCAAACCTGACAGAA 1929
DB 1741 GGCTCCACAGCAGCAGCAGCAAAACAAACGAGTGAATGTAGCCCTTCCAAACCTGACAGAA 1800
QY 1930 TGAGACCAACCGCAGCAGCAGCAGATCGGAGCAAAACCAAGACCAATCTGAGGATGAGAA 1989
DB 1801 TGAGACCAACCGCAGCAGCAGCAGATCGGAGCAAAACCAAGACCAATCTGAGGATGAGAA 1860
QY 1990 GTCTGGGAGCGCGCAGGAGCTCTGCCAGGCGCTGAGAACCCAGGCGCCGAGGAGGG 2049
DB 1861 GTCTGGGAGCGCGCAGGAGCTCTGCCAGGCGCTGAGAACCCAGGCGCCGAGGAGGG 1920
QY 2050 GCGGGAAGTCAAGCAGGTTTCCAGAACCAACGAGCCCGGCTCCCGCCCGCCAGGGC 2109
DB 1921 GCGGGAAGTCAAGCAGGTTTCCAGAACCAACGAGCCCGGCTCCCGCCCGCCAGGGC 1980
QY 2110 TTCTGCAAGCTTCAAGCAGCAGCTTCCATCCACTCCGATCTCTCTGAACTCCCAAG 2169
DB 1981 TTCTGCAAGCTTCAAGCAGCAGCTTCCATCCACTCCGATCTCTCTGAACTCCCAAG 2040
QY 2170 CGCTATCCCTTTAGTTGAATTAACATAGTGAACGTTTCAAAGCAACCAAAATGCAC 2229
DB 2041 CGCTATCCCTTTAGTTGAATTAACATAGTGAACGTTTCAAAGCAACCAAAATGCAC 2100
QY 2230 ACCCTTTTCTGTGGCAATCGCTCTGTACATGTGTGATCATATATAGAAAGGAGATG 2289
DB 2101 ACCCTTTTCTGTGGCAATCGCTCTGTACATGTGTGATCATATATAGAAAGGAGATG 2160
QY 2290 TTAGATATGTGGCTGTGGTTTACACAGGTCCTCGCAGCGGTATATATTTTAGAAAT 2349
DB 2161 TTAGATATGTGGCTGTGGTTTACACAGGTCCTCGCAGCGGTATATATTTTAGAAAT 2220
QY 2350 AATATATCAATAAATCAACTCAATCAATTTTAAATCAATTTTAAATTTTCTTTT 2409
DB 2221 AATATATCAATAAATCAACTCAATTTTAAATTTTAAATTTTCTTTTCTTT 2280
QY 2410 TTAAGAGAAAGCAGGCTTTTCTAGACTTTTAAAGATTAAGTCTTTGGGAGGTCTCAGG 2469
DB 2281 TTAAGAGAAAGCAGGCTTTTCTAGACTTTTAAAGATTAAGTCTTTGGGAGGTCTCAGG 2340
QY 2470 TGTAGAGAGGAGCTTTTGGAGCCACCAAAATTCACCCAGAGGAAATCTCGTCGGA 2529
DB 2341 TGTAGAGAGGAGCTTTTGGAGCCACCAAAATTCACCCAGAGGAAATCTCGTCGGA 2400
QY 2530 AGGACACTCAACGAGCTTCTGGATCACTGTGTATGTCAACAGAGGGAATACCGTCTCT 2589
DB 2401 AGGACACTCAACGAGCTTCTGGATCACTGTGTATGTCAACAGAGGGAATACCGTCTCT 2460
QY 2590 TGAAGAGAAACTCTGTCACTCTCATGCTGTAGCTCATACACCAATTTCTTTTTC 2649
DB 2461 TGAAGAGAAACTCTGTCACTCTCATGCTGTAGCTCATACACCAATTTCTTTTTC 2520
QY 2650 TTCAAGGTTTAAACGTTTGTGATATCTGATATATCTCTCTCTCTCTCTCTCTCT 2709
DB 2521 TTCAAGGTTTAAACGTTTGTGATATCTGATATATCTCTCTCTCTCTCTCTCTCT 2580
QY 2710 TATCTCTCCCTCTCCCTCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2769
DB 2581 TATCTCTCCCTCTCCCTCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2640
QY 2770 CCCTCCATCTCAATCCCGTATCTAGCAACCCCGCCCGCCCGCCCGCCCGCCCGCC 2829
DB 2641 CCCTCCATCTCAATCCCGTATCTAGCAACCCCGCCCGCCCGCCCGCCCGCCCGCC 2700
QY 2830 GTATCATCATCAACAAAGGAAACAAAGAGGAAACACACAAACCCAGGCTCTCACTTACACTT 2889

Db 2701 GTATCACATCACAAAAGGAAACAAAAGCGAAACACACAAACAGCCTCAACTTACACTT 2760
QY 2890 GGTACTCAAAAGACACAGAGTCAATGTTACTTTCCTAGCGTTTGGAGAGGAAACA 2949
Db 2761 GGTACTCAAAAGACACAGAGTCAATGTTACTTTCCTAGCGTTTGGAGAGGAAACA 2820
QY 2950 GGAACCCACCAACCAACCAATCAACCAACCAAGAAAGAAATTCACATGAAAGATGT 3009
Db 2821 GGAACCCACCAACCAACCAATCAACCAACCAAGAAAGAAATTCACATGAAAGATGT 2880
QY 3010 ATTTGTCTTTTGGATTTTGGTATAGCCATCAATATTCAGCAAAATGATTCCTTTC 3069
Db 2881 ATTTGTCTTTTGGATTTTGGTATAGCCATCAATATTCAGCAAAATGATTCCTTTC 2940
QY 3070 TTTAAAAAATAAATGTTGGAGAAAGTAGAATTTACCAAGTTTGTGGCCAGGCGTT 3129
Db 2941 TTTAAAAAATAAATGTTGGAGAAAGTAGAATTTACCAAGTTTGTGGCCAGGCGTT 3000
QY 3130 AAATTCACAGATTTTTTAAACGAGAAACACACAGAAAGCTACCTCAGGTGTTTTTA 3189
Db 3001 AAATTCACAGATTTTTTAAACGAGAAACACACAGAAAGCTACCTCAGGTGTTTTTA 3060
QY 3190 CTTAGACCTTGTCTTGTGTTTCCCTTAGAGATTTTGTAAAGCTGATGTTGAGCAT 3249
Db 3061 CTTAGACCTTGTCTTGTGTTTCCCTTAGAGATTTTGTAAAGCTGATGTTGAGCAT 3120
QY 3250 TTTTATTTTATTTTAAATAAATGATTTGGAAAAAATAGATATCACTGCCAGCTG 3309
Db 3121 TTTTATTTTATTTTAAATAAATGATTTGGAAAAAATAGATATCACTGCCAGCTG 3180
QY 3310 GAGAAAGTGACAGTCCAAAGTGTGCAACAGCTGTTCTGAATTTCTTCGCTAGCCAGAA 3369
Db 3181 GAGAAAGTGACAGTCCAAAGTGTGCAACAGCTGTTCTGAATTTCTTCGCTAGCCAGAA 3240
QY 3370 CNATATGCGCTCTTTTGGACAAACCTTGAATAATGTTTATTT 3412
Db 3241 CNATATGCGCTCTTTTGGACAAACCTTGAATAATGTTTATTT 3283

RESULT 5

AAS70981
ID AAS70981 standard; cDNA; 2010 BP.
AC AAS70981;
DT 13-FEB-2002 (first entry)
DE DNA encoding novel human diagnostic protein #6785.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX OS
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US008631.
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; A5G06794.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX Claim 1; SEQ ID NO 6785; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 2010 BP; 565 A; 552 C; 532 G; 361 T; 0 U; 0 Other;

Query Match 57.9%; Score 1977; DB 5; Length 2010;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1977; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GGGGAAGAGACGGATGATGAACAAGCTTTTACATCGGAAACCTGAGCCCGCGTCACCGC 116
Db 34 GGGGAAGAGACGGATGATGAACAAGCTTTTACATCGGAAACCTGAGCCCGCGTCACCGC 93
QY 117 CGAGACCTCCGGCAGCTCTTTGGGACAGGAAGCTGCCCTGGGGACAGAGTCTCTGCT 176
Db 94 CGAGACCTCCGGCAGCTCTTTGGGACAGGAAGCTGCCCTGGGGACAGAGTCTCTGCT 153
QY 177 GAAGTCCGGCTACGCTTTCTGTGACTACCCGACACAGAACTGGGCACTCCGCGCATCGA 236
Db 154 GAAGTCCGGCTACGCTTTCTGTGACTACCCGACACAGAACTGGGCACTCCGCGCATCGA 213
QY 237 GACCTCTCGGTTAAAGTGAATTCATCGGAAATCATGGAAGTTGATTAATCACTAGTCTC 296
Db 214 GACCTCTCGGTTAAAGTGAATTCATCGGAAATCATGGAAGTTGATTAATCACTAGTCTC 273
QY 297 TAAAAAGCTAAGGACAGGAAATTCAGATTCGAAACATCCCTCCTCACTCGAGTGGGA 356
Db 274 TAAAAAGCTAAGGACAGGAAATTCAGATTCGAAACATCCCTCCTCACTCGAGTGGGA 333
QY 357 GGTGTGGATGGACTTTTGGCTCAATATGGGACAGTGGAGAAATGTGGAACAAAGTCAAC 416
Db 334 GGTGTGGATGGACTTTTGGCTCAATATGGGACAGTGGAGAAATGTGGAACAAAGTCAAC 393
QY 417 AGACACAGAAACCGCGTTTGTCAACGTCAATATGCAACAGAGAAAGCAAAATAGC 476
Db 394 AGACACAGAAACCGCGTTTGTCAACGTCAATATGCAACAGAGAAAGCAAAATAGC 453
QY 477 CATGGAGAAGCTAAGCGGGCATCAGTTTGTAGAACTACTCTTCAAGATTTCTTACATCCC 536
Db 454 CATGGAGAAGCTAAGCGGGCATCAGTTTGTAGAACTACTCTTCAAGATTTCTTACATCCC 513
QY 537 GGATGAAGAGTGTAGCTCCCTTCGCCCTTCAGCGAGCCGAGCGTGGGGACCACTTTC 596
Db 514 GGATGAAGAGTGTAGCTCCCTTCGCCCTTCAGCGAGCCGAGCGTGGGGACCACTTTC 573
QY 597 CCGGAGCAAGGCCACCGCCCTCGGGGCATCTTCTCAGGCCAGACAGATTCGATTTCCCGCT 656
Db 574 CCGGAGCAAGGCCACCGCCCTCGGGGCATCTTCTCAGGCCAGACAGATTCGATTTCCCGCT 633

QY 657 GCGGATCTTGTCCTCCACCCAGTTTGTGTGTCATCATCGGAAAGAGGGCTTGACCAT 716
Db 634 GCGGATCTTGTCCTCCACCCAGTTTGTGTGTCATCATCGGAAAGAGGGCTTGACCAT 693
QY 717 AAGAAACATCACTAAGCAGACCCAGTCCCGGGTAGATATCCATAGAAAAAGAACTCTGG 776
Db 694 AAGAAACATCACTAAGCAGACCCAGTCCCGGGTAGATATCCATAGAAAAAGAACTCTGG 753
QY 777 AGCTGAGAGAAGCTGTCACTCCATCCATGCCACCCAGAGGGGACTTCTGAAGCATGCCG 836
Db 754 AGCTGAGAGAAGCTGTCACTCCATCCATGCCACCCAGAGGGGACTTCTGAAGCATGCCG 813
QY 837 CATGATCTTGAATCATGTCAGAAAGAGGCAGATGAGACCAACTAGCCGAAAGATTCC 896
Db 814 CATGATCTTGAATCATGTCAGAAAGAGGCAGATGAGACCAACTAGCCGAAAGATTCC 873
QY 897 TCTGAAAAATCTTGCGCACACAATGGCTTGTGGAAGACTGATTGGAAAAAGAGGCAGAA 956
Db 874 TCTGAAAAATCTTGCGCACACAATGGCTTGTGGAAGACTGATTGGAAAAAGAGGCAGAA 933
QY 957 TTTGAAAAAATTTGAAATGAACAGGACCAAGATAACAAATCTCATCTTTGACAGGATTT 1016
Db 934 TTTGAAAAAATTTGAAATGAACAGGACCAAGATAACAAATCTCATCTTTGACAGGATTT 993
QY 1017 GAGCATATACAAACCGGAAAGAACCATCACTGTGAAGGGCACAGTTGAGGCCTGTGCCAG 1076
Db 994 GAGCATATACAAACCGGAAAGAACCATCACTGTGAAGGGCACAGTTGAGGCCTGTGCCAG 1053
QY 1077 TGCTGAGATAGATATGAAAGAGCTGCGTGAAGCCCTTTGAAAAATGATATGCTGGCTGT 1136
Db 1054 TGCTGAGATAGATATGAAAGAGCTGCGTGAAGCCCTTTGAAAAATGATATGCTGGCTGT 1113
QY 1137 TAACCAACAAAGCCAAATCTGATCCAGGGTTGAACCTCAGCGCACTTGGCATCTTTTCAAC 1196
Db 1114 TAACCAACAAAGCCAAATCTGATCCAGGGTTGAACCTCAGCGCACTTGGCATCTTTTCAAC 1173
QY 1197 AGGACTGTGCGTGTATCTCCACAGCAGGCGCCCGCGAGCTCCCGCGCTGCCCTTA 1256
Db 1174 AGGACTGTGCGTGTATCTCCACAGCAGGCGCCCGCGAGCTCCCGCGCTGCCCTTA 1233
QY 1257 CCACCCCTTCACTACCCACTCCGGATATCTTCCAGAGCTGTACCCCATCACAGTTTGG 1316
Db 1234 CCACCCCTTCACTACCCACTCCGGATATCTTCCAGAGCTGTACCCCATCACAGTTTGG 1293
QY 1317 CCGGTTCCCGCATCATCACTCTTATCCAGAGCAGGAGATTGTGAATCTCTTCATCCCAAC 1376
Db 1294 CCGGTTCCCGCATCATCACTCTTATCCAGAGCAGGAGATTGTGAATCTCTTCATCCCAAC 1353
QY 1377 CCAGGCTGTGGGCCCATCATCGGAGAGAGGGGGCACATCAACAGCTGGCGAGATT 1436
Db 1354 CCAGGCTGTGGGCCCATCATCGGAGAGAGGGGGCACATCAACAGCTGGCGAGATT 1413
QY 1437 CGCCGGAGCCTCTATCAAGATTGCCCTTCCGGAAGGCCAGACCTCAGCGAAAGAGTGGT 1496
Db 1414 CGCCGGAGCCTCTATCAAGATTGCCCTTCCGGAAGGCCAGACCTCAGCGAAAGAGTGGT 1473
QY 1497 CATCATCCGGGCCACCGGAAGCCCGATTCAAGGCCCAGGACGATCTTTGGGAAACT 1556
Db 1474 CATCATCCGGGCCACCGGAAGCCCGATTCAAGGCCCAGGAGCGATCTTTGGGAAACT 1533
QY 1557 GAAAGAGAAACTCTTTTAAACCCCAAGAGAGAGTGAAGCTGGAAGCGCATATCAGAGT 1616
Db 1534 GAAAGAGAAACTCTTTTAAACCCCAAGAGAGAGTGAAGCTGGAAGCGCATATCAGAGT 1593
QY 1617 GCCCTCTTCCACAGCTGGCCGGTGAATTTGGCAAAAGTGGCAAGACCGTGAACGAACTGCA 1676
Db 1594 GCCCTCTTCCACAGCTGGCCGGTGAATTTGGCAAAAGTGGCAAGACCGTGAACGAACTGCA 1653
QY 1677 GAACTTAAACAGTGCAGAGTCACTGTCCTCGTGACCAAAAGCCAGATGAAATAGGA 1736
Db 1654 GAACTTAAACAGTGCAGAGTCACTGTCCTCGTGACCAAAAGCCAGATGAAATAGGA 1713
QY 1737 AGTGATCTCAGAAATATCGGGCACTCTTTGCTAGCCAGACTGCACAGCGCAAGATCAG 1796

Db 1714 AGTGATCGTCAGAAATATCGGCACATCTTTGCTAGCCAGACTGCACAGCGCAAGATCAG 1773
QY 1797 GGAAATTTGTACAACAGGTGAAGCAGCAGGAGCAGAAATACCCCTCAGGGAGTCGCCTCACA 1856
Db 1774 GGAAATTTGTACAACAGGTGAAGCAGCAGGAGCAGAAATACCCCTCAGGGAGTCGCCTCACA 1833
QY 1857 GGCACGAAGTGAAGTCCCAAGGACCAAGCAAAACAACGGATGAATGTAGCCCTTCCA 1916
Db 1834 GGCACGAAGTGAAGTCCCAAGGACCAAGCAAAACAACGGATGAATGTAGCCCTTCCA 1893
QY 1917 ACACCTTGACAGATGAGACCAACGACGACGACGATCGGAGCAAAACCAAGACCATC 1976
Db 1894 ACACCTTGACAGATGAGACCAACGACGACGACGATCGGAGCAAAACCAAGACCATC 1953
QY 1977 TGAGGAATGAGAGTCTGCGGAGCGCCAGGGAGCTTGCAGGGCCCTGAGAACCC 2033
Db 1954 TGAGGAATGAGAGTCTGCGGAGCGCCAGGGAGCTTGCAGGGCCCTGAGAACCC 2010

RESULT 6
AAS70982
ID AAS70982 standard; cDNA; 3667 BP.
XX
AC AAS70982;
XX DT 13-FEB-2002 (first entry)
XX DNA encoding novel human diagnostic protein #6786.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX WO200175067-A2.
XX PN 11-OCT-2001.
XX PD 30-MAR-2001; 2001WO-US008631.
XX PF 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; ABG06795.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 6786; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 3667 BP; 1081 A; 882 C; 846 G; 858 T; 0 U; 0 Other;

Query Match 53.0%; Score 1810; DB 5; Length 3667;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2100; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

1270 ACCCACTCGGATATCTTCCAGCTGTACCCCATCACAGTTTGGCCGTTCCGGCAT 1329
 1507 ACCCACTCGGATATCTTCCAGCTGTACCCCATCACAGTTTGGCCGTTCCGGCAT 1566
 1330 CATCATCTTATCCAGAGCAGGAGATTGTGAATCTTTCATCCCAACCCAGGCTGTGGC 1389
 1567 CATCATCTTATCCAGAGCAGGAGATTGTGAATCTTTCATCCCAACCCAGGCTGTGGC 1626
 1390 GCATCATCGGAAGAGGGGCGACACATCAACAGCTGGCGAGATTGGCCGAGCCTCT 1449
 1627 GCATCATCGGAAGAGGGGCGACACATCAACAGCTGGCGAGATTGGCCGAGCCTCT 1686
 1450 ATCAAGATTGCCCTCGCGAAGGCGCAGACGTCAGCGAAAGGATGGTCATCATACCGGG 1509
 1687 ATCAAGATTGCCCTCGCGAAGGCGCAGACGTCAGCGAAAGGATGGTCATCATACCGGG 1746
 1510 CCACCGAAGCCAGTCTCAAGGCCCAGGAGCGGATCTTTGGGAATCTGAAGAGAAAC 1569
 1747 CCACCGAAGCCAGTCTCAAGGCCCAGGAGCGGATCTTTGGGAATCTGAAGAGAAAC 1806
 1570 TTTCTTTAAACCAAGAGAGTGAAGCTGGAAGCGCATATCAGAGTCCCTCTTCCACA 1629
 1807 TTTCTTTAAACCAAGAGAGTGAAGCTGGAAGCGCATATCAGAGTCCCTCTTCCACA 1866
 1630 GCTGCGCGGGTGAATGGCAAGGTCGCAAGCCGTGAACGACACTCAGAACTTAAACAGT 1689
 1867 GCTGCGCGGGTGAATGGCAAGGTCGCAAGCCGTGAACGACACTCAGAACTTAAACAGT 1926
 1690 GCAGAGTCATCGTCCCTCGTGACCAACGCCAGATGAATGAGAAATGATCGTACA 1749
 1927 GCAGAGTCATCGTCCCTCGTGACCAACGCCAGATGAATGAGAAATGATCGTACA 1986
 1750 ATTATCGGCACTCTTTGTAGCCAGACTGCACAGCGCAAGATCAGGAAATGTACAA 1809
 1987 ATTATCGGCACTCTTTGTAGCCAGACTGCACAGCGCAAGATCAGGAAATGTACAA 2046
 1810 CAGGTGAAGCAGCAGGAGCAAAATACCTCAGGAGTCGCTCAGCGCAGCAAGTGA 1869
 2047 CAGGTGAAGCAGCAGGAGCAAAATACCTCAGGAGTCGCTCAGCGCAGCAAGTGA 2106
 1870 GGCTCCCAAGCAGCAGCAGCAAAACCAAGATGAATGATGAGCCTTCCAAACCTGACGAA 1929
 2107 GGCTCCCAAGCAGCAGCAGCAAAACCAAGATGAATGATGAGCCTTCCAAACCTGACGAA 2166
 1930 TGAGACCAAAACGACGACGAGATCGGAGCAAAACCAAGACCATCTGAGGAATGAGAA 1989
 2167 TGAGACCAAAACGACGACGAGATCGGAGCAAAACCAAGACCATCTGAGGAATGAGAA 2226
 1990 GTCTCGGAGCGCGCAGGAGCTCTGCGAGGCGCTTGAGAACCCAGGCGCGAGGAGG 2049
 2227 GTCTCGGAGCGCGCAGGAGCTCTGCGAGGCGCTTGAGAACCCAGGCGCGAGGAGG 2286
 2050 GCGGGAGGTCAGCCAGGTTGCGAGAACCAAGCAGCCCGCTCCGCGCCCGCAGGCG 2109
 2287 GCGGGAGGTCAGCCAGGTTGCGAGAACCAAGCAGCCCGCTCCGCGCCCGCAGGCG 2346
 2110 TTCTGCGAGGTCAGCCATCCATCTCAGATTCATCTGATCTCTCTGAACTCCACGA 2169
 2347 TTCTGCGAGGTCAGCCATCCATCTCAGATTCATCTGATCTCTCTGAACTCCACGA 2406

2170 CGCTATCCCTTTTGTAGTTGAACTAAACATAGCTGAACTGTTCAAAGCCAGCAAAATGCAC 2229
 2407 CGCTATCCCTTTTGTAGTTGAACTAAACATAGCTGAACTGTTCAAAGCCAGCAAAATGCAC 2466
 2230 ACCCTTTTCTGTGGCAAAATCGTCTGTACATGTTGTACATATTTAGAAAGGAGATG 2289
 2467 ACCCTTTTCTGTGGCAAAATCGTCTGTACATGTTGTACATATTTAGAAAGGAGATG 2526
 2290 TTAAGATATGTGGCCTGTGGGTTTACACAGGGTGCCTGAGCGGTAATATATTTTGAAT 2349
 2527 TTAAGATATGTGGCCTGTGGGTTTACACAGGGTGCCTGAGCGGTAATATATTTTGAAT 2586
 2350 AATATATCAAAATAACTCAACTAACTCCAAATTTTAAATCAATTTATTAATTTTCTTT 2409
 2587 AATATATCAAAATAACTCAACTAACTCCAAATTTTAAATCAATTTATTAATTTTCTTT 2646
 2410 TTAAGAGAAACGAGGCTTTTCTAGACTTTAAAGAAATAAAGTCTTTTGGAGGTCTCAAG 2469
 2647 TTAAGAGAAACGAGGCTTTTCTAGACTTTAAAGAAATAAAGTCTTTTGGAGGTCTCAAG 2706
 2470 TGTAGAGAGGAGCTTTGAGGCGCACCGGCACAAATTTACCCAGAGGAAATCTCTCGGA 2529
 2707 TGTAGAGAGGAGCTTTGAGGCGCACCGGCACAAATTTACCCAGAGGAAATCTCTCGGA 2766
 2530 AGGACACTCACGGCAGTTTCTGATCACCTGTGTATGTCAACAGAGGATACCGTCTCT 2589
 2767 AGGACACTCACGGCAGTTTCTGATCACCTGTGTATGTCAACAGAGGATACCGTCTCT 2826
 2590 TGAAGAGGAAACTCTGTCACTCTCATGCTGTCTAGCTCATACACCCATTTCTCTTTGC 2649
 2827 TGAAGAGGAAACTCTGTCACTCTCATGCTGTCTAGCTCATACACCCATTTCTCTTTGC 2886
 2650 TTCACAGTTTAAACTGGTTTTTGTATGTATGTATATATTTCTCTCTCTCTCTCTCT 2709
 2887 TTCACAGTTTAAACTGGTTTTTGTATGTATGTATATATTTCTCTCTCTCTCTCTCT 2946
 2710 TATCTCTCCCTCCCTCCCTCCCTCTCTCTCATCTCCATCTCTTTTGAATTTCTCTAT 2769
 2947 TATCTCTCCCTCCCTCCCTCCCTCTCTCTCATCTCCATCTCTTTTGAATTTCTCTAT 3006
 2770 CCCTCCATCTCAATCCCGTATCTAGCAACCCCTCCCTCCCTCCCTCCCTCCCTCTCTGA 2829
 3007 CCCTCCATCTCAATCCCGTATCTAGCAACCCCTCCCTCCCTCCCTCCCTCCCTCTCTGA 3065
 2830 GTATCATATCACACAAAGGAAACAAAGCGAAACACACAAACCCAGCTCACTTACACTT 2889
 3066 GTATCATATCACACAAAGGAAACAAAGCGAAACACACAAACCCAGCTCACTTACACTT 3125
 2890 GGTTACTCAAAAGAACAGAGTCAATGGTACTCTGTAGCGTTTGTGGAAGGAGAAACA 2949
 3126 GGTTACTCAAAAGAACAGAGTCAATGGTACTCTGTAGCGTTTGTGGAAGGAGAAACA 3185
 2950 GGAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAGT 3009
 3186 GGAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAGT 3245
 3010 ATTCTGTCTTTTGTGATTTGGTATTAAGCCATCAATATTCAGCAAAATGATTTCTTTTC 3069
 3246 ATTCTGTCTTTTGTGATTTGGTATTAAGCCATCAATATTCAGCAAAATGATTTCTTTTC 3305
 3070 TTT-AAAAAATAAAGTGTGAGGAAAGTAGAAATTTTACCAAGTTTGTGCGCCAGGCGGT 3128
 3306 TTTAAAAAATAAAGTGTGAGGAAAGTAGAAATTTTACCAAGTTTGTGCGCCAGGCGGT 3365
 3129 TAAATTCACAGATTTTAAACGAGAAACACACACAGAAAGCTACTCTAGGTTTCTTT 3188
 3366 TAAATTCACAGATTTTAAACGAGAAACACACACAGAAAGCTACTCTAGGTTTCTTT 3425
 3189 ACCTCAGCCTCTCTCTGTGTTTCTCTAGAGATTTGTAAAGCTGATGATTTGGAGCA 3248
 3426 ACCTCAGCCTCTCTGTGTTTCTCTAGAGATTTGTAAAGCTGATGATTTGGAGCA 3485
 3249 TTTTATTTTAAATAAATAAAGTGTGAGGAAATAAAGATATCAACTGCCAGCT 3308

Db	3486	TTTTTTTATTTTTTATATAAATGAGTTGGAAAAAATAAGATATCAACTCCAGCCT	3545
Qy	3309	GGAGAAGGTGACAGTCCAAAGTGTGCAACAGCTGTTCTGAATTGTTCTTCCGCTAGCCAAAGA	3368
Db	3546	GGAGAAGGTGACAGTCCAAAGTGTGCAACAGCTGTTCTGAATTGTTCTTCCGCTAGCCAAAGA	3605
Qy	3369	ACC 3371	
Db	3606	ACC 3608	
RESULT 7			
ABS76442			
ID ABS76442 standard; cDNA; 3667 BP.			
XX	ABS76442;		
XX	11-DEC-2002 (first entry)		
DT	XX		
XX	cDNA encoding human ovarian cancer marker M452.		
DE	Human; ovarian cancer; marker; cancer; familial history; brain disorder;		
XX	central nervous system disorder; bacterial meningitis; viral meningitis;		
KW	Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;		
KW	brain herniation; inflammation; encephalitis; testicular disorder;		
KW	nontuberculous granulomatous orchitis; connective tissue disorder;		
KW	heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;		
KW	histological type; carcinogenic; ovarian cancer marker; gene; ss.		
XX	Homo sapiens.		
OS	XX		
XX	WO200271928-A2.		
PN	19-SEP-2002.		
PD	XX		
XX	14-MAR-2002; 2002WO-US007826.		
PF	14-MAR-2001; 2001US-0276025P.		
XX	14-MAR-2001; 2001US-0276026P.		
PR	10-AUG-2001; 2001US-0311732P.		
PR	19-SEP-2001; 2001US-0323580P.		
PR	26-SEP-2001; 2001US-0324967P.		
PR	26-SEP-2001; 2001US-0325102P.		
PR	26-SEP-2001; 2001US-0325149P.		
XX	(MILL-) MILLENNIUM PHARM INC.		
PA	XX		
XX	Monahan JE, Gannavarapu M, Hoersch S, Kamatkar S, Kovatis SG;		
PI	Meyers RE, Morrissey MP, Olandt PJ, Sen A, Vieby PO, Mills GB;		
PI	Bast RC, Lu K, Schmandt RE, Zhao X, Glatt K;		
PI	WPI; 2002-723277/78.		
DR	P-PSDB; ABG96346.		
XX	Assessing whether a patient is afflicted with ovarian cancer, useful in		
PT	assessing the stage or progression of the disease, comprises comparing		
PT	the expression level of a cancer marker in a sample from a patient and		
PT	from a non cancer patient.		
XX	Disclosure; Page 262-263; 481pp; English.		
XX	The present invention relates to a new method for assessing whether a		
CC	patient is afflicted with ovarian cancer. The method involves comparing		
CC	the expression level of a marker in a patient sample and the normal level		
CC	of expression of the marker in a control non-ovarian cancer sample, where		
CC	the marker is selected from 363 cancer markers described in the		
CC	specification. The method of the invention is useful in diagnosing or		
CC	characterising cancer, in detecting the presence of cancer as early as		
CC	possible, and the recurrence of ovarian cancer. The method may also be		
CC	particular use with patients having an enhanced risk of developing		
CC	ovarian cancer (e.g. patients having a familial history of ovarian		
CC	cancer). The cancer markers may be used in the management and treatment		

QY 2050 GCGGGAAGGTGAGCCAGGTTTGCCAGAACACCGAGCCCGCTCCCGCCCCCAGG3C 2109
Db 2287 GCGGGAAGGTGAGCCAGGTTTGCCAGAACACCGAGCCCGCTCCCGCCCCCAGG3C 2346
QY 2110 TTCTGAGGCTTCAGCCATCCACTTCAACATCCACTCGGATCTCTCTGAACTCCACGA 2169
Db 2347 TTCTGAGGCTTCAGCCATCCACTTCAACATCCACTCGGATCTCTCTGAACTCCACGA 2406
QY 2170 GCGTATCCCTTTAGTCTGAACCTAAACATAGTGTGAAAGCGGAAATGCA 2229
Db 2407 GCGTATCCCTTTAGTCTGAACCTAAACATAGTGTGAAAGCGGAAATGCA 2466
QY 2230 ACCCTTTTCTGTGGCAATCGTCTCTGTACATGTGTGTACATATTAGAAAGGGAAGTG 2289
Db 2467 ACCCTTTTCTGTGGCAATCGTCTCTGTACATGTGTGTACATATTAGAAAGGGAAGTG 2526
QY 2290 TTAAGATATGTGGCTGTGGGTTACAGAGGTGCTGACGGGTAAATATTATTTAGAAAT 2349
Db 2527 TTAAGATATGTGGCTGTGGGTTACAGAGGTGCTGACGGGTAAATATTATTTAGAAAT 2586
QY 2350 AATATATCAATACTCAACTAACTCAATTTTAAATCAATTAATTTTCTTTTCTTT 2409
Db 2587 AATATATCAATACTCAACTAACTCAATTTTAAATCAATTAATTTTCTTTTCTTT 2646
QY 2410 TTAAGAGAAAGCAGGCTTTTCTAGACTTTTAAAGATAAAGTCTTTGGGAGGTCTCACGG 2469
Db 2647 TTAAGAGAAAGCAGGCTTTTCTAGACTTTTAAAGATAAAGTCTTTGGGAGGTCTCACGG 2706
QY 2470 TGTAGAGAGGACTTTGAGGCCACCGGCACAAAATTCACCGAGGAAATCTGTCGA 2529
Db 2707 TGTAGAGAGGACTTTGAGGCCACCGGCACAAAATTCACCGAGGAAATCTGTCGA 2766
QY 2530 AGGACACTCAGGCAAGTCTGTGATCACTGTGTATGTCAACAGAGGATACCGTCTCT 2589
Db 2767 AGGACACTCAGGCAAGTCTGTGATCACTGTGTATGTCAACAGAGGATACCGTCTCT 2826
QY 2590 TGAAGAGAAACTGTCTCACTCTCATGCTGTCTAGCTCTACACCCATTTCTTTTGC 2649
Db 2827 TGAAGAGAAACTGTCTCACTCTCATGCTGTCTAGCTCTACACCCATTTCTTTTGC 2886
QY 2650 TTCACAGGTTTAACTGGTTTTTGTGATCTGTATATATATCTCTCTCTCTCTCT 2709
Db 2887 TTCACAGGTTTAACTGGTTTTTGTGATCTGTATATATATCTCTCTCTCTCTCT 2946
QY 2710 TATCTCTCCCTCCCTCCCTCCCTCTCTCTCACTCTCACTCTCTCTCTCTCTCTCAT 2769
Db 2947 TATCTCTCCCTCCCTCCCTCCCTCTCTCTCACTCTCACTCTCTCTCTCTCTCTCAT 3006
QY 2770 CCTTCATCTCAATCCCGTATCTAGCAACCCCGCCCCCCCCCAGGCAAGCGTCTCTGA 2829
Db 3007 CCTTCATCTCAATCCCGTATCTAGCAACCCCGCCCCCCCCCAGGCAAGCGTCTCTGA 3065
QY 2830 GTATCATATCACAAAGGACAAAGGAAACACACAAACCCAGCCTCACTTACACTT 2889
Db 3066 GTATCATATCACAAAGGACAAAGGAAACACACAAACCCAGCCTCACTTACACTT 3125
QY 2890 GGTACTCAAAAGAACAGAGTCAATGTGTCTGTCTAGCGTTTGTGAAGAGGAAACA 2949
Db 3126 GGTACTCAAAAGAACAGAGTCAATGTGTCTGTCTAGCGTTTGTGAAGAGGAAACA 3185
QY 2950 GGAACCCACCAACCAACCAATCAACCAACAAAGAAATTTCCCAATGAAGATGT 3009
Db 3186 GGAACCCACCAACCAACCAATCAACCAACAAAGAAATTTCCCAATGAAGATGT 3245
QY 3010 ATTTTGTCTTTTGCATTTTGGTGTATAGGCAATATATTCAGCAAAATGATTCCTTTC 3069
Db 3246 ATTTTGTCTTTTGCATTTTGGTGTATAGGCAATATATTCAGCAAAATGATTCCTTTC 3305
QY 3070 TTT-AAAAAATAATGTGAGGAAAGTAGAAATTTTACCAAGTTTGTGGCCCGAGGGGT 3128
Db 3306 TTTAAAAAATAATGTGAGGAAAGTAGAAATTTTACCAAGTTTGTGGCCCGAGGGGT 3365

QY 3129 TAAATTCACAGATTTTTTAAACGAGAAAAACACACAGAGAAAGCTACCTCAGGTGTTTT 3188
Db 3366 TAAATTCACAGATTTTTTAAACGAGAAAAACACACAGAGAAAGCTACCTCAGGTGTTTT 3425
QY 3189 ACCTCAGCACCTTGCTCTTGTGTTTCCCTTAGAGATTTTGTAAAGCTGATAGTTGGAGCA 3248
Db 3426 ACCTCAGCACCTTGCTCTTGTGTTTCCCTTAGAGATTTTGTAAAGCTGATAGTTGGAGCA 3485
QY 3249 TTTTATTTTATTTTAAATAAAATGAGTTGGAAAAAATAAGATATCAATCGCAGCCT 3308
Db 3486 TTTTATTTTATTTTAAATAAAATGAGTTGGAAAAAATAAGATATCAATCGCAGCCT 3545
QY 3309 GGAGAAAGTGACAGTCCCAAGTGTGCAACAGCTGTTCTCAATTTGCTTCGCTAGCAAGA 3368
Db 3546 GGAGAAAGTGACAGTCCCAAGTGTGCAACAGCTGTTCTGAATTTGCTTCGCTAGCAAGA 3605
QY 3369 ACC 3371
Db 3606 ACC 3608

RESULT 8
AAS26153
ID AAS26153 standard; cDNA; 2290 BP.
XX
AC AAS26153;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human cDNA encoding a novel secreted protein, Seq ID 332.
XX
KW Human; immunosuppressive; antiarthritic; ss; antirheumatic; cytostatic;
KW cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;
KW antibacterial; virucide; fungicide; ophthalmological; vulnerary;
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; skin ageing; food additive;
KW preservative; antiproliferative.
XX
OS Homo sapiens.
XX
PN WO200155322-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001341.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 11-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0217496P.
PR 26-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.

CC capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors and other nutritional components. The present CC sequence encodes a novel secreted protein of the invention. Note: The CC sequence data for this patent did not form part of the printed	Query Match Best Local Similarity 43.8%; Score 1496; DB 4; Length 2290; Matches 1966; Conservative 0; Mismatches 1; Indels 4; Gaps 3;
Qy 1387	GGCGCCATCATCGGGAAGAAGGGGGCCACATCAAAAGCTGGCGAGATTCGCCGGAGCC 1446
Db 36	GGCGCCATCATCGGGAAGAAGGGGGCCACATCAAAAGCTGGCGAGATTCGCCGGAGCC 95
Qy 1447	TCTATCAAGATTGCCCTTGCGGAAGGCCAGACTCAGCGAAAGGATGGTCAATCATCACC 1506
Db 96	TCTATCAAGATTGCCCTTGCGGAAGGCCAGACTCAGCGAAAGGATGGTCAATCATCACC 155
Qy 1507	GGGCGACGGGAAGCCCAAGTTCAAAGCCCCAGGGACGGATCTTTGGGAACCTGAAGAGGAA 1566
Db 156	GGGCGACGGGAAGCCCAAGTTCAAAGCCCCAGGGACGGATCTTTGGGAACCTGAAGAGGAA 215
Qy 1567	AACTTCTTTAAACCCAAAGAAAGTAGTGAAGCTGGAAGCGCATATCAGAGTGCCTCTTCC 1626
Db 216	AACTTCTTTAAACCCAAAGAAAGTAGTGAAGCTGGAAGCGCATATCAGAGTGCCTCTTCC 2175
Qy 1627	ACAGCTGGCCGGGTGATTGGCAAAAGGTGGCAAGACCGTGAAACGAACTTCAGAACTTAAAC 1686
Db 276	ACAGCTGGCCGGGTGATTGGCAAAAGGTGGCAAGACCGTGAAACGAACTTCAGAACTTAAAC 335
Qy 1687	AGTCAGAAGTCATCGTGCCCTCGTGCACCAAGCCAGATGAAATGAGGAAGTGATCGTC 1746
Db 336	AGTCAGAAGTCATCGTGCCCTCGTGCACCAAGCCAGATGAAATGAGGAAGTGATCGTC 395
Qy 1747	AGAATTATCGGGCACTTCTTTGTAGCCAGATGSCACAGCGCAAGATCAGGGAATTTGTA 1806
Db 396	AGAATTATCGGGCACTTCTTTGTAGCCAGATGSCACAGCGCAAGATCAGGGAATTTGTA 455
Qy 1807	CAACAGGTGAGCAGCAGGAGCAGAAATACCTCAGGAGTGGCTCACAGCGCAGCAAG 1866
Db 456	CAACAGGTGAGCAGCAGGAGCAGAAATACCTCAGGAGTGGCTCACAGCGCAGCAAG 515
Qy 1867	TGAGGCTCCCAACAGGCCACAGCAAAAACACGGATGAATGTAGCCCTTCCAAACCTTGACA 1926
Db 516	TGAGGCTCCCAACAGGCCACAGCAAAAACACGGATGAATGTAGCCCTTCCAAACCTTGACA 575
Qy 1927	GAATGAGACAAACGCGCAGCCAGATCGGGAGCAAAACCAAGACCATCTGAGGAATGA 1986
Db 576	GAATGAGACAAACGCGCAGCCAGATCGGGAGCAAAACCAAGACCATCTGAGGAATGA 635
Qy 1987	GAAGTCTGCGAGGCGCGCAGGACTCTGCGAGGCGCTGAGAACCCAGGGCGCAGGA 2046
Db 636	GAAGTCTGCGAGGCGCGCAGGACTCTGCGAGGCGCTGAGAACCCAGGGCGCAGGA 695
Qy 2047	GGGGCGGGGAAGGTACGCCAGGTTTGGCAAAACACCGAGCCCGCTCCCGCCCCCAG 2106
Db 696	GGGGCGGGGAAGGTACGCCAGGTTTGGCAAAACACCGAGCCCGCTCCCGCCCCCAG 755
Qy 2107	GGCTTCTGAGGCTTACGCCATCACTTACCATCCAGTGGATCTCTCTGAACTCCCA 2166
Db 756	GGCTTCTGAGGCTTACGCCATCACTTACCATCCAGTGGATCTCTCTGAACTCCCA 815
Qy 2167	CGAGCGTATCCCTTTTAGTTGAACTACATAGTGAACTGTTCAAGCCCAAGCAAAATG 2226
Db 816	CGAGCGTATCCCTTTTAGTTGAACTACATAGTGAACTGTTCAAGCCCAAGCAAAATG 875
Qy 2227	CACACCCCTTTTCTGTGGCAAACTCGTCTCTGTATCATGTGTGTACATATTAGAAAGGGAAG 2286
Db 876	CACACCCCTTTTCTGTGGCAAACTCGTCTCTGTATCATGTGTGTACATATTAGAAAGGGAAG 935
Qy 2287	ATGTTAAGATATGTGGCTGTGGTTACACAGGGTGCCTGCAGCGGTATATATTTTGA 2346
Db 936	ATGTTAAGATATGTGGCTGTGGTTACACAGGGTGCCTGCAGCGGTATATATTTTGA 995

QY	2347	AATTAATATATACAAATACTCAACTCAATTTTAAATTAATTAATTTTTC	2406
Db	996	AATTAATATATCAAAATACTCAACTCAATTTTAAATTAATTAATTTTTC	1055
QY	2407	TTTTTAAAGAGAAAGCAGGCTTTCTAGACTTTAAAGAAATAAAGTCTTTGGGAGGTCTCA	2466
Db	1056	TTTTTAAAGAGAAAGCAGGCTTTCTAGACTTTAAAGACTTTTGGGAGGTCTCA	1115
QY	2467	CGGTGTAGAGAGAGCTTTGAGGCCACCCCGCAAAAATTCACCGAGGGGAAATCTCGTC	2526
Db	1116	CGGTGTAGAGAGAGCTTTGAGGCCACCCCGCAAAAATTCACCGAGGGGAAATCTCGTC	1175
QY	2527	GGAAGGACACTCACGGCAGTTCTGGATCAGCTGTGTATGTCAACAGAGGGATACCGTCT	2586
Db	1176	GGAAGGACACTCACGGCAGTTCTGGATCAGCTGTGTATGTCAACAGAGGGATACCGTCT	1235
QY	2587	CTTTGAAGAGGAAACTCTGTCACCTCCTATGCTCTGTCTAGCTCATACGCCATTTCTCTT	2646
Db	1236	CTTTGAAGAGGAAACTCTGTCACCTCCTATGCTCTGTCTAGCTCATACGCCATTTCTCTT	1295
QY	2647	TGCTTCACAGGTTTTAAACTGGTTTTTGTGACTACTGCTATATAATTTCTGTCTCTCTCT	2706
Db	1296	TGCTTCACAGGTTTTAAACTGGTTTTTGTGACTACTGCTATATAATTTCTGTCTCTCTCT	1355
QY	2707	GTTTATCTCTCCCTCCCTCCCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2766
Db	1356	G-TTATCTCTCCCTCCCTCCCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1414
QY	2767	CATCCCTCCATCTCAATCCCGTATCTAGGCA--CCCCCCCCCCCCCAGGCAAGCAGTGC	2824
Db	1415	CATCCCTCCATCTCAATCCCGTATCTAGGCACCCCCCCCCCCCCAGGCAAGCAGTGC	1474
QY	2825	TCTGAGTATCAGATCACAACAAAAGGAAACAAAAGGAAACACACAAACCGACCTCAACTTA	2884
Db	1475	TCTGAGTATCAGATCACAACAAAAGGAAACAAAAGGAAACACACAAACCGACCTCAACTTA	1534
QY	2885	CACCTGGTTACTCAAAGAAACAAGACTCAATGTTACTTGTCTAGCGTTTTCGAAAGGA	2944
Db	1535	CACCTGGTTACTCAAAGAAACAAGACTCAATGTTACTTGTCTAGCGTTTTCGAAAGGA	1594
QY	2945	AAACAGGAAACCCACCAACCAACCAATCAACCAACCAACCAACCAACCAACCAATGAAAG	3004
Db	1595	AAACAGGAAACCCACCAACCAACCAATCAACCAACCAACCAACCAACCAACCAATGAAAG	1654
QY	3005	AATGTAATTTGCTCTTTTGTGATTTGCTGTATAAGCCATCAATATTCAGCAAAATCATTC	3064
Db	1655	AATGTAATTTGCTCTTTTGTGATTTGCTGTATAAGCCATCAATATTCAGCAAAATCATTC	1714
QY	3065	CTTTCTTTT-AAAAAAAAATGTGGAGGAAAGTAGAAATTTTACCAAGGTTGTGGCCACG	3123
Db	1715	CTTTCTTTTAAAAAAAAAATGTGGAGGAAAGTAGAAATTTTACCAAGGTTGTGGCCACG	1774
QY	3124	GGCGTTAAATTCACAGATTTTTTAAACGAGAAACACACAGAGAGAGCTACTCAGGTG	3183
Db	1775	GGCGTTAAATTCACAGATTTTTTAAACGAGAAACACACAGAGAGAGCTACTCAGGTG	1834
QY	3184	TTTTTACCTCAGCACCTTGTCTGTGTTTTCCCTTAGAGATTTTGTAAAGCTGATAGTTG	3243
Db	1835	TTTTTACCTCAGCACCTTGTCTGTGTTTTCCCTTAGAGATTTTGTAAAGCTGATAGTTG	1894
QY	3244	GAGCATTTTTTTATTTTTTAAATAAAATAGGTTGGAATAAAATAGATATCAACTGCC	3303
Db	1895	GAGCATTTTTTTATTTTTTAAATAAAATAGGTTGGAATAAAATAGATATCAACTGCC	1954
QY	3304	AGCCTGGAGAGGTGACGTCGAAAGTGTGCAACAGCTGTCTGAAATGTCT	3354
Db	1955	AGCCTGGAGAGGTGACGTCGAAAGTGTGCAACAGCTGTCTGAAATGTCT	2005

RESULT 9
ABX73494
ID ABX
XX

AC ABX73494;
XX 18-MAR-2003 (first entry)
XX Human novel polynucleotide #322.
DE
XX
XX Human; gene; ds; neural disorder; immune system disorder; renal disorder;
KW muscular disorder; respiratory disease; reproductive disorder;
KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
KW hyperproliferative disorder; inflammatory disease; allergic reaction;
KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
KW haemostatic; antiarteriosclerotic.
XX
OS Homo sapiens.
XX
PN US2002132753-A1.
XX
XX 19-SEP-2002.
XX
XX 17-JAN-2001; 2001US-00764864.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214886P.
PR 07-JUL-2000; 2000US-0216847P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234937P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
XX 08-DEC-2000; 2000US-0251869P.

PA (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
XX (BARA/) BARASH S C.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2003-147444/14.
DR P-PSDB; ABUS5234.
XX
XX New polypeptides and nucleic acids, useful in gene therapy for treating,
PT inhibiting or preventing e.g. neural, immune system, muscular,
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
PT renal disorders.
XX
PS Claim 1; SEQ ID NO 332; 402pp; English.
XX
XX The invention relates to human novel polypeptides and their associated
CC polynucleotides. The polypeptides and polynucleotides are useful in gene
CC therapy for treating, inhibiting or preventing neural disorders, immune
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
CC infarction) and cancerous diseases. Sequences ABX73173-ABX74167 represent
CC human novel polynucleotides of the invention
XX
SQ Sequence 2290 BP; 707 A; 561 C; 491 G; 531 T; 0 U; 0 Other;

Query Match 43.8%; Score 1496; DB 7; Length 2290;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1966; Conservative 0; Mismatches 1; Indels 4; Gaps 3;

QY 1387 GCGCCATCATCGGGGGAAGAGGGGGGCACATCAAAAGCTGGCGAGATTGGCGGAGCC 1446
DB 36 GCGCCATCATCGGGGGAAGAGGGGGGCACATCAAAAGCTGGCGAGATTGGCGGAGCC 95

QY 1447 TCTATCAAGATTGCCCCCTCGGAAGGCCAGAGCTGAGCAAGAGTGTCTCATCATCACC 1506
DB 96 TCTATCAAGATTGCCCCCTCGGAAGGCCAGAGCTGAGCAAGAGTGTCTCATCATCACC 155

QY 1507 GGGCCACCGAAGCCCGAGTTCAAGGGCCAGGGACGGATCTTTGGGAACTGAAAGAGGAA 1566
DB 156 GGGCCACCGAAGCCCGAGTTCAAGGGCCAGGGACGGATCTTTGGGAACTGAAAGAGGAA 215

QY 1567 AACTTCTTTAAACCCCAAGAAAGTGAAGCTGGAAGCGCATATCAGAGTGCCCTCTTCC 1626
DB 216 AACTTCTTTAAACCCCAAGAAAGTGAAGCTGGAAGCGCATATCAGAGTGCCCTCTTCC 275

QY 1627 ACAGCTGGCCGGGTGATTGGCAAGGTGCAACACCGTGAACGAACCTGCAGAACTTAACC 1686
DB 276 ACAGCTGGCCGGGTGATTGGCAAGGTGCAACACCGTGAACGAACCTGCAGAACTTAACC 335

QY 1687 AGTGCAGAGTATCGTGCCTCGTGCACCAACCCAGATGAAATGAGGAAGTATCGTC 1746
DB 336 AGTGCAGAGTATCGTGCCTCGTGCACCAACCCAGATGAAATGAGGAAGTATCGTC 395

QY 1747 AGAATTATCGGGCACTTTCTTTGCTAGCCAGACTGCACAGCGCAAGATCAGGGAATTTGA 1806
DB 396 AGAATTATCGGGCACTTTCTTTGCTAGCCAGACTGCACAGCGCAAGATCAGGGAATTTGA 455

QY 1807 CAACAGGTGAAGCAGGAGGAGCAATACCTCAGGGAGTCCCTCAGAGCCAGCAAG 1866
DB 456 CAACAGGTGAAGCAGGAGGAGCAATACCTCAGGGAGTCCCTCAGAGCCAGCAAG 515

QY 1867 TGAGGCTCCCAACAGGACCAAGCAAAACACGAGTGAATGTAGCCCTTCCAAACCTGACA 1926
DB 516 TGAGGCTCCCAACAGGACCAAGCAAAACACGAGTGAATGTAGCCCTTCCAAACCTGACA 575

QY	1927	GAATGAGACCAAA	CGCAGCC	CAGCAGAT	TCGGAGAGCAAA	CAAGACCAAT	CTGAGGAATGA	1986
Db								
QY	576	GAATGAGACCAAA	CGCAGCC	CAGCAGAT	TCGGAGAGCAAA	CAAGACCAAT	CTGAGGAATGA	635
Db								
QY	1987	GAAGTCTCGGAGG	CGGCC	AGGACT	CTGCCGAGCGCCT	GAGAACCC	CCAGGGCCGAGGA	2046
Db								
QY	636	GAAGTCTCGGAGG	CGGCC	AGGACT	CTGCCGAGCGCCT	GAGAACCC	CCAGGGCCGAGGA	695
Db								
QY	2047	GGGGCGGGAAAG	TGAC	CGGTTTG	CGCAGAAC	CACGAGCCCGCCT	CCGCGCCGCCAG	2106
Db								
QY	696	GGGGCGGGAAAG	TGAC	CGGTTTG	CGCAGAAC	CACGAGCCCGCCT	CCGCGCCGCCAG	755
Db								
QY	2107	GGCTCTGACAGCT	TAC	GCATTC	CACATTC	CACTCGAT	TCTCTGAACTCCCA	2166
Db								
QY	756	GGCTCTGACAGCT	TAC	GCATTC	CACATTC	CACTCGAT	TCTCTGAACTCCCA	815
Db								
QY	2167	CGAGCTATCCCT	TTT	TAGTTG	AACTAA	CATAGGTG	AAAGCTGTTCAAAAGCCAAAGT	2226
Db								
QY	816	CGAGCTATCCCT	TTT	TAGTTG	AACTAA	CATAGGTG	AAAGCTGTTCAAAAGCCAAAGT	875
Db								
QY	2227	CACACCCCTTT	TTT	TGTG	CAAAATCG	TCTGTAC	ATGTTGATCATATTTAGAAAGGGAAG	2286
Db								
QY	876	CACACCCCTTT	TTT	TGTG	CAAAATCG	TCTGTAC	ATGTTGATCATATTTAGAAAGGGAAG	935
Db								
QY	2287	ATGTTAAGATAT	GTG	GGCTGT	GCGTTAC	AGGTGCT	GACGGGTAAATATATTTTAGA	2346
Db								
QY	936	ATGTTAAGATAT	GTG	GGCTGT	GCGTTAC	AGGTGCT	GACGGGTAAATATATTTTAGA	995
Db								
QY	2347	AATAATATAT	CAAAATA	ACTCAACT	CAACTTTC	CAATTTTAA	CTCAATTAATTAATTTTTC	2406
Db								
QY	996	AATAATATAT	CAAAATA	ACTCAACT	CAACTTTC	CAATTTTAA	CTCAATTAATTAATTTTTC	1055
Db								
QY	2407	TTTTTAAAGAA	AGAC	GGCTTT	TCTAG	ACTTTAA	AGAATAAAGTCTTTGGGAGGTCTCA	2466
Db								
QY	1056	TTTTTAAAGAA	AGAC	GGCTTT	TCTAG	ACTTTAA	AGAATAAAGTCTTTGGGAGGTCTCA	1115
Db								
QY	2467	CGGTGTAGAG	GAGCTTT	GAGGCC	ACCGCA	CAAAAT	TCCACGAGGGAATCTCGTC	2526
Db								
QY	1116	CGGTGTAGAG	GAGCTTT	GAGGCC	ACCGCA	CAAAAT	TCCACGAGGGAATCTCGTC	1175
Db								
QY	2527	GGAAGGAC	ACT	CAC	GCAGTCT	GGATCA	CTCTGTATGTCAACAGAAAGGATACCGTCT	2586
Db								
QY	1176	GGAAGGAC	ACT	CAC	GCAGTCT	GGATCA	CTCTGTATGTCAACAGAAAGGATACCGTCT	1235
Db								
QY	2587	CGTTGAGAG	GAAT	ACTCT	GTAC	TCTCT	CATCCGTCTAGCTCATACACCAATTTCTCTT	2646
Db								
QY	1236	CGTTGAGAG	GAAT	ACTCT	GTAC	TCTCT	CATCCGTCTAGCTCATACACCAATTTCTCTT	1295
Db								
QY	2647	TGCTTCAC	AGGTTTT	TAACT	TGGTTTT	TGCAT	TACTGCTATATAATTTCTCTCTCT	2706
Db								
QY	1296	TGCTTCAC	AGGTTTT	TAACT	TGGTTTT	TGCAT	TACTGCTATATAATTTCTCTCTCT	1355
Db								
QY	2707	GTTTATCT	CTCCCT	CCCT	CCCT	TCTTCT	CTCATCTCCATTTCTTTGAATTTCT	2766
Db								
QY	1356	G-TATATCT	CTCCCT	CCCT	CCCT	TCTTCT	CTCATCTCCATTTCTTTGAATTTCT	1414
Db								
QY	2767	CATCCCTCC	CAAT	CCCGTAT	CTAC	GAC	CCCCCCCCCCCCCAGGCAAAAGCAGTGC	2824
Db								
QY	1415	CATCCCTCC	CAAT	CCCGTAT	CTAC	GAC	CCCCCCCCCCCCCAGGCAAAAGCAGTGC	1474
Db								
QY	2825	TCTGAGTAT	CACAT	CAC	AAAAAGGA	CAAAAGGGA	CAACAAACCGAGCTCACTTA	2884
Db								
QY	1475	TCTGAGTAT	CACAT	CAC	AAAAAGGA	CAAAAGGGA	CAACAAACCGAGCTCACTTA	1534
Db								
QY	2885	CACCTTG	TTTACT	CAAAAGAA	CAAGTCA	ATGGT	TACTTTGTCTTAGCGTTTGTGAAGAGGA	2944
Db								
QY	1535	CACCTTG	TTTACT	CAAAAGAA	CAAGTCA	ATGGT	TACTTTGTCTTAGCGTTTGTGAAGAGGA	1594
Db								
QY	2945	AAACAGGAA	CCCCAC	CAAAAC	CAAACT	CAAC	CAAAACAAAGAAAAAATCCCAATGAAAG	3004
Db								
QY	1595	AAACAGGAA	CCCCAC	CAAAAC	CAAACT	CAAC	CAAAACAAAGAAAAAATCCCAATGAAAG	1654
Db								

QY	3005	AATGTAATTTGTCCTTTTTCGATTTTGGTGATTAAGCAATCAATATTACAGCAAAATGATTC	3066
Db	1655	AATGTAATTTGTCCTTTTTCGATTTTGGTGATTAAGCAATCAATATTACAGCAAAATGATTC	1714
QY	3065	CTTTCTTTT-AAAAAATAAATGTTGGAGAAAGTAGAAATTTTACCAAGGTTGTTGGCCCG	3123
Db	1715	CTTTCTTTT-AAAAAATAAATGTTGGAGAAAGTAGAAATTTTACCAAGGTTGTTGGCCCG	1774
QY	3124	GGCGTTAAATTCACAGATTTTTTTTAAACGAGAAAAACACACAGAGAAGAGCTACCTCAGGTG	3183
Db	1775	GGCGTTAAATTCACAGATTTTTTTTAAACGAGAAAAACACACAGAGAAGAGCTACCTCAGGTG	1834
QY	3184	TTTTTACCTCAGCACCTTGCTCTGTGTTTCCCTTAGAGATTTTGTAAAGCTGATAGTTG	3243
Db	1835	TTTTTACCTCAGCACCTTGCTCTGTGTTTCCCTTAGAGATTTTGTAAAGCTGATAGTTG	1894
QY	3244	GAGCAATTTTTTTTATTTTTTAAATAAAATGAGTTGGAAAAAATAAGATATCAACTGCC	3303
Db	1895	GAGCAATTTTTTTTATTTTTTAAATAAAATGAGTTGGAAAAAATAAGATATCAACTGCC	3303
QY	3304	AGCCTGAGAAAGTGACAGTCCAAAGTGTGCAACAGCTGTTCTGAATTGTCT	3354
Db	1955	AGCCTGAGAAAGTGACAGTCCAAAGTGTGCAACAGCTGTTCTGAATTGTCT	2005
RESULT 10			
AAA02565			
ID	AAA02565 standard; cDNA; 710 BP.		
XX	AC AAA02565;		
XX	AC AAA02565;		
DT	19-MAY-2000 (first entry)		
XX	Human colon cancer cell line polynucleotide sequence SEQ ID NO:2556.		
DE	Human; colon cancer; tumour; diagnosis; gene expression product; probe;		
KW	detection; cancerous state; metastasis; identification; breast cancer;		
KW	oestrogen receptor-positive breast cancer; therapy;		
KW	oestrogen receptor-negative breast cancer; lung cancer; ss.		
OS	Homo sapiens.		
XX	WO9958675-A2.		
XX	18-NOV-1999.		
PF	13-MAY-1999; 99WO-US010602.		
XX	14-MAY-1998; 98US-0085426P.		
PR	15-MAY-1998; 98US-0085537P.		
PR	15-MAY-1998; 98US-0085696P.		
PR	21-OCT-1998; 98US-0105234P.		
PR	27-OCT-1998; 98US-0105877P.		
XX	(CHIR) CHIRON CORP.		
PA	(HYSE-) HYSEQ INC.		
XX	Williams LT, Escobedo J, Inniss MA, Garcia PD, Sudduth-Klinger J;		
PI	Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;		
PI	Lamson G, Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;		
PI	Leshkowitz D, Xia D, Garcia V, Jones LW, Stache-Crain B;		
XX	WPI; 2000-126369/11.		
DR	Polynucleotide library used to determine cancerous states of mammalian		
XX	cells.		
PT	Claim 1; Page 1034; 1097pp; English.		
XX	AAA00010 to AAA02716 represent polynucleotides isolated from cDNA		
CC	libraries constructed from human colon cancer cell lines. The present		
CC	invention also describes a method of detecting differentially expressed		
CC	genes correlated with a cancerous state of a mammalian cell comprising		

CC detecting at least one differentially expressed gene product in a test
CC sample derived from a cell suspected of being cancerous, where detection
CC of the differentially expressed gene product is correlated with a
CC cancerous state of the cell from which the test sample was derived. The
CC polynucleotide sequences can be used in a method for detecting
CC differentially expressed genes correlated with a cancerous state of a
CC mammalian cell. The polynucleotides can also be used as probes for
CC detecting and mapping related genes. They can be used in diagnosis and
CC prognosis of diseases and disorders (e.g. identification of pre-
CC metastatic or metastatic cancerous states, stages of cancer, or
CC responsiveness of cancer to therapy). This is particularly for breast
CC cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-
CC negative breast cancer, lung cancer, and colon cancer
XX
SQ Sequence 710 BP; 199 A; 174 C; 180 G; 152 T; 0 U; 5 Other;
Query Match 16.6%; Score 567; DB 3; Length 710;
Best Local Similarity 100.0%; Pred. No. 2.2e-243;
Matches 567; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 534 CCGGATGAAGAGGTGAGCTCCCTTCGCCCCCTCAGCGAGCCAGCGTGGGACCACTC 593
Db 25 CCGGATGAAGAGGTGAGCTCCCTTCGCCCCCTCAGCGAGCCAGCGTGGGACCACTC 84
QY 594 TTCCCGGAGCAAGCGCCCTCGGGGCACTTCTCAGGCCAGACAGATTGATTTCCTCC 653
Db 85 TTCCCGGAGCAAGCGCCCTCGGGGCACTTCTCAGGCCAGACAGATTGATTTCCTCC 144
QY 654 GCTGCGGATCCTGGTCCCAACCCAGTTTGTGTGCTCATCGGAAAGGAGGCTTGAC 713
Db 145 GCTGCGGATCCTGGTCCCAACCCAGTTTGTGTGCTCATCGGAAAGGAGGCTTGAC 204
QY 714 CATAAAGAACATCACTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAAGAACTC 773
Db 205 CATAAAGAACATCACTAAGCAGACCCAGTCCCGGTAGATATCCATAGAAAAGAACTC 264
QY 774 TGAGCTGCAGAGAGCTGTCCATCCATGCCACCCAGAGGGGACTTCTGAGCATG 833
Db 265 TGGAGCTGCAGAGAGCTGTCCATCCATGCCACCCAGAGGGGACTTCTGAGCATG 324
QY 834 CCGCATGATCTTGAATCATGCAAGAGGAGGAGTGAACCAACTAGCCGAGAGAT 893
Db 325 CCGCATGATCTTGAATCATGCAAGAGGAGGAGTGAACCAACTAGCCGAGAGAT 384
QY 894 TCCTCTGAAATCTTGGCACAATGCGCTTGGTGGGAAGTCAATGGAAGAAAGGCAG 953
Db 385 TCCTCTGAAATCTTGGCACAATGCGCTTGGTGGGAAGTCAATGGAAGAAAGGCAG 444
QY 954 AAATTTGAAGAAATTGAACATGAAACAGGACCCAGATACCAATCTCATCTTTGCAGGA 1013
Db 445 AAATTTGAAGAAATTGAACATGAAACAGGACCCAGATACCAATCTCATCTTTGCAGGA 504
QY 1014 TTTGAGCATATACAACCCGAAAGAACCACTCTGTGAAAGGACAGCTTGAGGCGCTGTC 1073
Db 505 TTTGAGCATATACAACCCGAAAGAACCACTCTGTGAAAGGACAGCTTGAGGCGCTGTC 564
QY 1074 CAGTGTGAGATAGAGATTATGAAGAA 1100
Db 565 CAGTGTGAGATAGAGATTATGAAGAA 591

RESULT 11
AAS26570
ID AAS26570 standard; cDNA; 1186 BP.
XX AC AAS26570;
XX DT 07-NOV-2001 (first entry)
XX Human cDNA encoding a novel secreted protein, seq ID 749.
DE Human; immunosuppressive; antiarthritic; ss; antirheumatic; cytostatic;
KW Human; vasotropic; cerebroprotective; nootropic; neuroprotective;
KW

KW antibacterial; virucide; fungicide; ophthalmological; vulnery;
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; skin ageing; food additive;
KW preservative; antiproliferative.
XX Homo sapiens.
XX WO200155322-A2.
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US001341.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUL-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 11-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225466P.
PR 14-AUG-2000; 2000US-0225467P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.

PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234224P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0250300P.
PR 01-DEC-2000; 2000US-0250316P.
PR 05-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0256719P.
PR 08-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.

PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX
PI Rosen CA, Barash SC, Ruben SM;
DR WPI; 2001-488783/53.
DR P-PSDB; AAU16583.
XX
XX
PT New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives.
XX
PS Claim 1; SEQ ID NO 749; 980pp; English.
XX
CC The invention relates to isolated nucleic acid molecules and their
CC encoded secreted proteins. The nucleic acids and proteins are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. Antibodies to the proteins can also be used in
CC alleviating symptoms associated with the disorders and in diagnostic
CC immunassays e.g. radioimmunoassays or enzyme linked immunosorbant assays
CC (ELISA). Disorders which are diagnosed or treated include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiodysplasia,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
CC and many other disorders listed in the specification. The polypeptides
CC can also be used to aid wound healing and epithelial cell proliferation,
CC to prevent skin aging due to sunburn, to maintain organs before,
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence encodes a novel secreted protein of the invention. Note: The
CC sequence data for this patent did not form part of the printed

Query Match 11.8%; Score 401; DB 4; Length 1186;

Best Local Similarity 99.4%; Pred. No. 5.1e-169;
Matches 671; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY	1864	AAGTGGGCTCCACAGGACACAGCAAAACAAAGGATGATGTAGCCCTTCCAAACCTG	1923
Db	512	AAGTGGGCTCCACAGGACACAGCAAAACAAAGGATGATGTAGCCCTTCCAAACCTG	571
QY	1924	ACAGATGAGACAAACGACAGCCAGATCGGAGCAAAACCAAGACCATCTGAGGAA	1983
Db	572	ACAGATGAGACAAACGACAGCCAGATCGGAGCAAAACCAAGACCATCTGAGGAA	631
QY	1984	TGAGAGTCTGCGAGGCGGCGGAGGACTCTGCGAGGCGCTGAGAACCCAGGGGCGGA	2043
Db	632	TGAGAGTCTGCGAGGCGGCGGAGGACTCTGCGAGGCGCTGAGAACCCAGGGGCGGA	691
QY	2044	GGAGGGGCGGGGAAGGTTCAGCCAGGTTTGCCAGAACCCAGGCGCCGCGCCGCC	2103
Db	692	GGAGGGGCGGGGAAGGTTCAGCCAGGTTTGCCAGAACCCAGGCGCCGCGCCGCC	751
QY	2104	CAGGGCTTCTGAGGCTTCAGGCAAT-CCACTTCACCATCCACTCGGATCTCTCTGAAC	2162
Db	752	CAGGGCTTCTGAGGCTTCAGGCAAT-CCACTTCACCATCCACTCGGATCTCTCTGAAC	811
QY	2163	CCCAACGAGTATCCCTTTTAGTTGAACATAAGTAGTGAACGTGTTCAAGCCAGCAA	2222
Db	812	CCCAACGAGTATCCCTTTTAGTTGAACATAAGTAGTGAACGTGTTCAAGCCAGCAA	871

QY 2223 AATGCACACCCCTTTTCTGTGGCAATCGTCTGTACATGTGTACATATTAGAAAG 2282
DB 872 AATGCACACCCCTTTTCTGTGGCAATCGTCTGTACATGTGTACATATTAGAAAG 931
QY 2283 GAAGATGTTAAGATATGTGGCTGTGGGTTACACAGGTCCTGACGGGTAATATATT 2342
DB 932 GAAGATGTTAAGATATGTGGCTGTGGGTTACACAGGTCCTGACGGGTAATATATT 991
QY 2343 TAGAATAATATATCAATAACTCAACTCAACTCAACTCAACTCAACTCAACTCAACT 2402
DB 992 TAGAATAATATATCAATAACTCAACTCAACTCAACTCAACTCAACTCAACTCAACT 1051
QY 2403 TTCTCTTTTAAAG 2462
DB 1052 TTCTCTTTTAAAG 1111
QY 2463 CTCACGGTGTAG 2522
DB 1112 CTCACGGTGTAG 1171
QY 2523 CGTCGGAAGACACT 2537
DB 1172 CGTCGGAAGACACT 1186
RESULT 12
ABX73911
ID ABX73911 standard; DNA; 1186 BP.
XX
AC ABX73911;
XX
DT 18-MAR-2003 (first entry)
XX
DE Human novel polynucleotide #739.
XX
DE Human; gene; ds; neural disorder; immune system disorder; renal disorder;
KW muscular disorder; respiratory disease; reproductive disorder;
KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
KW hyperproliferative disorder; inflammatory disease; allergic reaction;
KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
KW haemostatic; antiarteriosclerotic.
XX
OS Homo sapiens.
XX
PN US2002132753-A1.
XX
PD 19-SEP-2002.
XX
PF 17-JAN-2001; 2001US-00764864.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214896P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216809P.
PR 11-JUL-2000; 2000US-0217497P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX
PA (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
PI WPI; 2003-147444/14.
XX
DR P-PSDB; ABUS5651.
XX
PT New polypeptides and nucleic acids, useful in gene therapy for treating,
PT inhibiting or preventing e.g. neural, immune system, muscular,
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
PT renal disorders.
XX
PS Claim 1; SEQ ID NO 749; 402pp; English.
XX
CC The invention relates to human novel polypeptides and their associated
CC polynucleotides. The polypeptides and polynucleotides are useful in gene
CC therapy for treating, inhibiting or preventing neural disorders, immune
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
CC leukaemia), inflammatory diseases (e.g. septic shock, burstitis and
CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
CC infarction) and cancerous diseases. Sequences ABX73173-ABX74167 represent
CC human novel polynucleotides of the invention
XX
SQ Sequence 1186 BP; 338 A; 310 C; 302 G; 230 T; 0 U; 6 Other;
Query Match 11.8%; Score 401; DB 7; Length 1186;
Best Local Similarity 99.4%; Pred. No. 5.1e-169;
Matches 671; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
QY 1864 AGTGAGGCTCCACAGGACCCAGCAAAACACGATGATGTAGCCCTTCCACACTG 1923
DB 512 AGTGAGGCTCCACAGGACCCAGCAAAACACGATGATGTAGCCCTTCCACACTG 571
QY 1924 ACAGATGAGACCAACCGCAGCCAGATCGGGAGCAAAACCAAGACCATCTGAGGAA 1983
|||||

Db	572	ACAGATGAGACCAACGACGACCGACAGATCGGGAGCAACCAACCAAGACCATCTGAGAA	631	PR 24-FEB-2000; 2000US-01846644
QY	1984	TGAGAAAGTCTCGGAGGCGGCACAGGACTCTGCCGAGGCGCCCTGAGAAACCCACGAGGGCCGA	2043	PR 02-MAR-2000; 2000US-0186350P
Db	632	TGAGAAAGTCTCGGAGGCGGCACAGGACTCTGCCGAGGCGCCCTGAGAAACCCACGAGGGCCGA	691	PR 16-MAR-2000; 2000US-0189874P
QY	2044	GGAGGGGCGGGAGGTCAGCCAGGTTTGCAGAACCAACGAGCCCGCTCCGCCCCCC	2103	PR 17-MAR-2000; 2000US-0190076P
Db	692	GGAGGGGCGGGAGGTCAGCCAGGTTTGCAGAACCAACGAGCCCGCTCCGCCCCCC	751	PR 18-APR-2000; 2000US-0198123P
QY	2104	CAGGGCTTCTCGAGGCTTCAGCCAT-COACTTCACCATCCACTCGGATCTCTCTGAACT	2162	PR 19-MAY-2000; 2000US-0205515P
Db	752	CAGGGCTTCTCGAGGCTTCAGCCATTCAGCCATTCACCATTCCTCTCTGAACT	811	PR 27-JUN-2000; 2000US-0209467P
QY	2163	CCACGACGCTATCCCTTTTAGTTGAACTAACTAGGTGAACGTTTCAAAGCCAAAGCAA	2222	PR 30-JUN-2000; 2000US-0215135P
Db	812	CCACGACGCTATCCCTTTTAGTTGAACTAACTAGGTGAACGTTTCAAAGCCAAAGCAA	871	PR 07-JUL-2000; 2000US-0216647P
QY	2223	AATGCACACCCCTTTTCTGTGGCAATCGTCTCTGTACATGTGTGTACATATTAGAAAGG	2282	PR 14-AUG-2000; 2000US-0224518P
Db	872	AATGCACACCCCTTTTCTGTGGCAATCGTCTCTGTACATGTGTGTACATATTAGAAAGG	931	PR 14-AUG-2000; 2000US-0225213P
QY	2283	GAGATGTTAAGATATGTGGCTGTGGTTACACAGGGTCCCTGCAGCGTAAATATATTT	2342	PR 14-AUG-2000; 2000US-0225267P
Db	932	GAGATGTTAAGATATGTGGCTGTGGTTACACAGGGTCCCTGCAGCGTAAATATATTT	991	PR 14-AUG-2000; 2000US-0225270P
QY	2343	TAGAAATAATATCAATAACTCAACTCACTCCAAATTTTAAATCAATTAATTTT	2402	PR 14-AUG-2000; 2000US-0225547P
Db	992	TAGAAATAATATCAATAACTCAACTCACTCCAAATTTTAAATCAATTAATTTT	1051	PR 14-AUG-2000; 2000US-0225757P
QY	2403	TTTCTTTTAAAGAGAAAGCAGGCTTTTCTAGACTTTTAAAGAATAAAGTCTTTGGAGGT	2462	PR 14-AUG-2000; 2000US-0225759P
Db	1052	TTTCTTTTAAAGAGAAAGCAGGCTTTTCTAGACTTTTAAAGAATAAAGTCTTTGGAGGT	1111	PR 18-AUG-2000; 2000US-0226279P
QY	2463	CTCAGGTGTAGAGAGAGCTTTGAGCCACCCGCGCACAAATTCACCAGAGGAAATCT	2522	PR 22-AUG-2000; 2000US-0226681P
Db	1112	CTCAGGTGTAGAGAGAGCTTTGAGCCACCCGCGCACAAATTCACCAGAGGAAATCT	1171	PR 22-AUG-2000; 2000US-0226686P
QY	2523	CGTCGGAAGGACACT 2537		PR 23-AUG-2000; 2000US-0227009P
Db	1172	CGTCGGAAGGACACT 1186		PR 30-AUG-2000; 2000US-0228924P
RESULT 13				PR 01-SEP-2000; 2000US-0229287P
AAS26151				PR 01-SEP-2000; 2000US-0229343P
ID AAS26151 standard; cDNA; 822 BP.				PR 01-SEP-2000; 2000US-0229344P
XX				PR 01-SEP-2000; 2000US-0229345P
AC				PR 05-SEP-2000; 2000US-0229509P
AC				PR 05-SEP-2000; 2000US-0229513P
XX				PR 06-SEP-2000; 2000US-0230437P
XX				PR 06-SEP-2000; 2000US-0230438P
XX				PR 08-SEP-2000; 2000US-0231242P
XX				PR 08-SEP-2000; 2000US-0231243P
XX				PR 08-SEP-2000; 2000US-0231244P
XX				PR 08-SEP-2000; 2000US-0231413P
XX				PR 08-SEP-2000; 2000US-0231414P
XX				PR 08-SEP-2000; 2000US-0232080P
XX				PR 12-SEP-2000; 2000US-0232081P
XX				PR 12-SEP-2000; 2000US-0231968P
XX				PR 14-SEP-2000; 2000US-0232397P
XX				PR 14-SEP-2000; 2000US-0232398P
XX				PR 14-SEP-2000; 2000US-0232399P
XX				PR 14-SEP-2000; 2000US-0232399P
XX				PR 14-SEP-2000; 2000US-0232401P
XX				PR 14-SEP-2000; 2000US-0233063P
XX				PR 14-SEP-2000; 2000US-0233064P
XX				PR 14-SEP-2000; 2000US-0233065P
XX				PR 21-SEP-2000; 2000US-0234233P
XX				PR 21-SEP-2000; 2000US-0234274P
XX				PR 25-SEP-2000; 2000US-0234997P
XX				PR 25-SEP-2000; 2000US-0234998P
OS				PR 26-SEP-2000; 2000US-0235484P
XX				PR 27-SEP-2000; 2000US-0235834P
XX				PR 27-SEP-2000; 2000US-0235836P
XX				PR 29-SEP-2000; 2000US-0236327P
XX				PR 29-SEP-2000; 2000US-0236367P
XX				PR 29-SEP-2000; 2000US-0236367P
XX				PR 29-SEP-2000; 2000US-0236369P
XX				PR 29-SEP-2000; 2000US-0236370P
XX				PR 02-OCT-2000; 2000US-0236802P
XX				PR 02-OCT-2000; 2000US-0237037P
XX				PR 02-OCT-2000; 2000US-0237038P

Human cDNA encoding a novel secreted protein, Seq ID 330.

Human; immunosuppressive; antiarthritic; ss; antirheumatic; cytostatic; cardiac; vasotropic; cerebroprotective; nootropic; neuroprotective; antibacterial; virucide; fungicide; ophthalmological; vulnerary; secreted protein; rheumatoid arthritis; hyperproliferative disorder; cardiovascular disorder; cardiac arrest; cerebrovascular disorder; cerebral ischaemia; angiogenesis; nervous system disorder; Alzheimer's disease; infection; ocular disorder; corneal infection; wound healing; epithelial cell proliferation; skin ageing; food additive; preservative; antiproliferative.

Homo sapiens.

WO200155322-A2.

02-AUG-2001.

17-JAN-2001; 2001WO-US001341.

31-JAN-2000; 2000US-0179065P.

04-FEB-2000; 2000US-0180628P.

PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 17-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249284P.
PR 17-NOV-2000; 2000US-0249285P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0256719P.
PR 08-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-488783/53.
XX P-FSDB; AAU16164.
XX
XX New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives.
XX

PS Claim 1; SEQ ID NO 330; 980pp; English.
XX The invention relates to isolated nucleic acid molecules and their
CC encoded secreted proteins. The nucleic acids and proteins are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. Antibodies to the proteins can also be used in
CC alleviating symptoms associated with the disorders and in diagnostic
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbant assays
CC (ELISA). Disorders which are diagnosed or treated include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
CC and many other disorders listed in the specification. The polypeptides
CC can also be used to aid wound healing and epithelial cell proliferation,
CC to prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence encodes a novel secreted protein of the invention. Note: The
CC sequence data for this patent did not form part of the printed

Query Match 11.4%; Score 389; DB 4; Length 822;
Best Local Similarity 99.8%; Pred. No. 1.2e-163;
Matches 439; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 779 CTGCAGAGAGCCTGTCTCACCATCCATGCCACCCAGAGGGGACTTCTGAAGCATGCCGCA 838
Db 1 CTGCAGAGAGCCTGTCTCACCATCCATGCCACCCAGAGGGGACTTCTGAAGCATGCCGCA 60
Qy 839 TGATTTCTTGAATCATGCAGAAAGAGGCAGATGAGACCAAACTAGCCGAGAGATTCTCTC 898
Db 61 TGATTTCTTGAATCATGCAGAAAGAGGCAGATGAGACCAAACTAGCCGAGAGATTCTCTC 120
Qy 899 TGAATAATCTTGGCACACAATGGCTTGGTTGGAAGACTGTATTGGAAGAGAGGCGAGAAAT 958
Db 121 TGAATAATCTTGGCACACAATGGCTTGGTTGGAAGACTGTATTGGAAGAGAGGCGAGAAAT 180
Qy 959 TGAAGAAAATGAAACATGAAACAGGACCAAGATAACAATCTCATCTTTGCAGGATTGA 1018
Db 181 TGAAGAAAATGAAACATGAAACAGGACCAAGATAACAATCTCATCTTTGCAGGATTGA 240
Qy 1019 GCATATACAAACCCGGAAGAACCATCACTGTGAAGGGCACAGTTGAGGCCCTGTGCCAGTG 1078
Db 241 GCATATACAAACCCGGAAGAACCATCACTGTGAAGGGCACAGTTGAGGCCCTGTGCCAGTG 300
Qy 1079 CTGAGATAGAGATTATGAAGAAGCTGCGTGAGGCGCTTTGAAAATGATATGCTGCTGTGA 1138
Db 301 CTGAGATAGAGATTATGAAGAAGCTGCGTGAGGCGCTTTGAAAATGATATGCTGCTGTGA 360
Qy 1139 ACCAAGAGCAATCTCATCTCCAGGGTTGAACCTCAGGCACCTGGCATCTTTTCAACAG 1198
Db 361 ACCAAGAGCAATCTCATCTCCAGGGTTGAACCTCAGGCACCTGGCATCTTTTCAACAG 420
Qy 1199 GACTGTCCGTCTATCTCCA 1218
Db 421 GACTGTCCGTCTATCTCCA 440

RESULT 14
ABX73492
ID ABX73492 standard; DNA; 822 BP.
XX
AC ABX73492;
XX
DT 18-MAR-2003 (first entry)
XX
DE Human novel polynucleotide #320.

XX Human; gene; ds; neural disorder; immune system disorder; renal disorder;
KW muscular disorder; respiratory disease; reproductive disorder; disorder;
KW gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
KW hyperproliferative disorder; inflammatory disease; allergic reaction;
KW blood related disorder; cancer; immunosuppressive; antiinflammatory;
KW cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
KW haemostatic; antiarteriosclerotic.
XX
OS Homo sapiens.
XX
XX US2002132753-A1.
PN
XX 19-SEP-2002.
PD
XX
XX 17-JAN-2001; 2001US-00764864.
PF
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214886P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX
XX (ROSE/) ROSEN C A.
FA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
XX
PI Rosen CA, Ruben SM, Barash SC;

XX WPI: 2003-147444/14.
DR P-PSDB; ABUS5232.
XX
PT New polypeptides and nucleic acids, useful in gene therapy for treating,
PT inhibiting or preventing e.g. neural, immune system, muscular,
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
PT renal disorders.
XX
XX Claim 1; SEQ ID NO 330; 402pp; English.
PS
XX
CC The invention relates to human novel polypeptides and their associated
CC polynucleotides. The polypeptides and polynucleotides are useful in gene
CC therapy for treating, inhibiting or preventing neural disorders, immune
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
CC infarction) and cancerous diseases. Sequences ABX73173-ABX74167 represent
CC human novel polynucleotides of the invention
XX
SQ Sequence 822 BP; 234 A; 212 C; 201 G; 163 T; 0 U; 12 Other;

Query Match 11.4%; Score 389; DB 7; Length 822;
Best Local Similarity 99.8%; Pred. No. 1.2e-163;
Matches 439; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 779 CTGCAGAGAGCGCTGTACCATCCATGCCACCCAGAGGGGACTTCTGAAGCATGCCGA 838
Db 1 CTGCAGAGAGCGCTGTACCATCCATGCCACCCAGAGGGGACTTCTGAAGCATGCCGA 60
QY 839 TGATTCTTGAATCATGCAGAAAGAGGCAGATGAGACCAACTAGCCGAGAGATTCTTC 898
Db 61 TGATTCTTGAATCATGCAGAAAGAGGCAGATGAGACCAACTAGCCGAGAGATTCTTC 120
QY 899 TGAATAATCTTGGCACACAATCGCTTGGTTGGAAGACTGATTGGAAAAGAGCAAAATT 958
Db 121 TGAATAATCTTGGCACACAATCGCTTGGTTGGAAGACTGATTGGAAAAGAGCAAAATT 180
QY 959 TGAAGAAATTTGAACATGAACACAGGACCAAGATACCAATCTCATCTTTGAGGATTGA 1018
Db 181 TGAAGAAATTTGAACATGAACACAGGACCAAGATACCAATCTCATCTTTGAGGATTGA 240
QY 1019 GCATATACACCCGAAAGAACCATCACTGTGAAGGGCACAGTTGAGGCCCTGTGCCAGTG 1078
Db 241 GCATATACACCCGAAAGAACCATCACTGTGAAGGGCACAGTTGAGGCCCTGTGCCAGTG 300
QY 1079 CTGAGATAGAGATTATGAAGAGCTGCGTGAAGGCTTTGAAAAATCATATGCTGGCTGTTA 1138
Db 301 CTGAGATAGAGATTATGAAGAGCTGCGTGAAGGCTTTGAAAAATCATATGCTGGCTGTTA 360
QY 1139 ACCAACAGCAATCTGATCCAGGCTTGACCTCAGGCACTTGGCATCTTTTCAACAG 1198
Db 361 ACCAACAGCAATCTGATCCAGGCTTGAACTCCTCAGGCACTTGGCATCTTTTCAACAG 420
QY 1199 GACTGTCCGTGCTATCTCCA 1218
Db 421 GACTGTCCGTGCTATCTCCA 440

RESULT 15
AAC03267
ID AAC03267 standard; cDNA; 364 BP.
XX
AC AAC03267;
XX
DT 06-OCT-2000 (first entry)

Search completed: July 21, 2004, 17:06:27
Job time : 1245 secs

```
XX Human secreted protein 5' EST, SEQ ID NO: 3265.
DE
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX Homo sapiens.
OS
XX EP1033401-A2.
PN
XX 06-SEP-2000.
PD
XX 21-FEB-2000; 2000EP-00200610.
PF
XX 26-FEB-1999; 99US-0122487P.
PR
XX (GEST ) GENSET.
PA
XX Dumas Milne Edwards J, Duclert A, Giordano J;
PI
XX WPI; 2000-500381/45.
DR
XX P-PSDB; AAG03261.
DR
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
PT
XX Claim 1; SEQ ID NO 3265; 71pp + Sequence Listing; English.
PS
XX The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. An ORP has been identified within the
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC derived from 30 different tissues. EST sequences usually correspond
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors
CC
XX Sequence 364 BP; 84 A; 101 C; 118 G; 59 T; 0 U; 2 Other;
Qy Query Match 10.2%; Score 347; DB 3; Length 364;
Db Best Local Similarity 100.0%; Pred. No. 8.1e-145;
Matches 347; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 15 CGAGGAGCGCGGGTACCGGGCCGGGGAGCGCGGGCTCTCGGGGAAGAGACGGATGAT 74
Db 18 CGAGGAGCGCGGGTACCGGGCCGGGGAGCGCGGGCTCTCGGGGAAGAGACGGATGAT 77
Qy 75 GAACAAGCTTTACATCGGACCTGACCCCGCGGTACCGCCGACGACCTCCGGCAGCT 134
Db 78 GAACAAGCTTTACATCGGACCTGACCCCGCGGTACCGCCGACGACCTCCGGCAGCT 137
Qy 135 CTTTGGGACAGGAAGCTGCCCTCGGGGACAGGTCTCTGTAAGTCCGGCTACGCCTT 194
Db 138 CTTTGGGACAGGAAGCTGCCCTCGGGGACAGGTCTCTGTAAGTCCGGCTACGCCTT 197
Qy 195 CGTGGACTACCCGACAGAACTGGGCCATCCGGCCATCGAGACCCCTCTCGGGTAAAGT 254
Db 198 CGTGGACTACCCGACAGAACTGGGCCATCCGGCCATCGAGACCCCTCTCGGGTAAAGT 257
Qy 255 GGAATTGATCGGAAATCATGGAAGTTGATTACTCAGTCTCTAAAGCTAAGAGCAG 314
Db 258 GGAATTGATCGGAAATCATGGAAGTTGATTACTCAGTCTCTAAAGCTAAGAGCAG 317
Qy 315 GAAATTTCAGATTCGAAACATCCCTCCCTCAGTGGGAGGTGT 361
Db 318 GAAATTTCAGATTCGAAACATCCCTCCCTCAGTGGGAGGTGT 364
```

This Page Blank (uspto)